
 RELEASE
 (TM)

Release 3.1A John F. Collins, Bioinformatics Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

MParch_n n.e. n.a. database search, using Smith-Waterman algorithm

Run on: Sun Feb 14 21:30:34 1999; Maspar time 353.20 Seconds

Tabular output not generated. 921.792 Million cell updates/sec

Title: >US-08-946-869-9
 Description: (1-2393) from US08946869.seq
 Perfect Score: 2393
 N.A. Sequence: 1 AAGGCCCTGCGCGCTCTC.....AAAAAAGCGCGCGCGGT 2393
 Comp: TTCGGGTGGAGCGCGAGAG.....TTTTTCCGCGCGCGCGCA

Scoring table: TABLE default

Gap 6

Mmatch STD : Dbase 0; Query 0

Searched: 188442 seqs, 68026449 bases x 2

Post-processing: Minimum Match 0

Listing first 45 summaries

Database:

n:geneseq2
 1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
 8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
 14:part14 15:part15 16:part16 17:part17 18:part18
 19:part19 20:part20 21:part21 22:part22 23:part23
 24:part24 25:part25 26:part26 27:part27 28:part28
 29:part29 30:part30 31:part31 32:part32 33:part33
 34:part34 35:part35 36:part36 37:part37 38:part38
 39:part39 40:part40

Statistics: Mean 9.947; Variance 7.935; scale 1.253

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	116	4.8	116 22	T25991	Human gene signature	1.45e-41
2	48	2.0	91	Q51746	Oligonucleotide probe	6.92e-09
3	45	1.9	204	N81164	Base substituted E.co	1.44e-07
4	43	1.8	91	Q51746	Oligonucleotide probe	1.05e-06
5	42	1.8	204	N81164	Base substituted E.co	2.85e-06
6	40	1.7	114 12	T70468	Generic DNA sequence	2.03e-05
7	40	1.7	172 32	T76363	Human interleukin 8 a	2.03e-05
8	40	1.7	946 13	Q80219	Human NDF-alpha2b clo	2.03e-05
9	41	1.7	1098 13	Q80225	Rat NDF clone 4 DNA.	2.03e-05
10	40	1.7	1651 13	Q80218	Human NDF-alpha2b clo	2.03e-05
11	40	1.7	1734 35	T39792	Human SHP9 gene.	2.03e-05
12	40	1.7	3344 33	Q80228	Rat NDF clone 22 DNA.	2.03e-05
13	40	1.7	3451 38	V02308	Cell membrane proton-	2.03e-05

Result ID	Score	Query Match	Length	DB ID	Description	Pred. No.
14	38	1.6	114 12	Q70466	Generic DNA sequence	1.40e-04
15	39	1.6	178 32	T76405	Human endothelin-1 an	3.35e-05
16	38	1.6	369 38	V02169	Human secreted protei	1.40e-04
17	38	1.6	1425 23	T88086	3' portion of cDNA cl	1.40e-04
18	38	1.6	1425 23	T35230	Cytoplasmic antiprote	1.40e-04
19	39	1.6	2335 13	Q80216	Human prointra-alpha2b	3.35e-05
20	38	1.6	2400 1	Q05055	Placenta-specific CD	1.40e-04
21	38	1.6	2862 14	Q84612	Mouse ATR2 receptor cl	1.40e-04
22	36	1.5	56 9	Q52732	Sequence of oligo nuc	9.40e-04
23	36	1.5	114 12	Q70467	Generic DNA sequence	9.40e-04
24	36	1.5	114 12	Q70469	Generic DNA sequence	9.40e-04
25	36	1.5	114 12	Q70470	Generic DNA sequence	9.40e-04
26	36	1.5	114 12	Q70467	Generic DNA sequence	9.40e-04
27	36	1.5	162 32	T76307	Human RANTES antisens	9.40e-04
28	36	1.5	201 38	V00418	3' fragment of clone	9.40e-04
29	36	1.5	607 38	V02138	Human secreted protei	9.40e-04
30	36	1.5	607 39	T97397	Human secreted protei	9.40e-04
31	36	1.5	607 39	T88059	Partial cDNA clone en	9.40e-04
32	36	1.5	701 32	T77001	Mouse PLA2s cDNA (B6a	9.40e-04
33	36	1.5	810 32	T48835	Mouse PLA2s cDNA (W11	9.40e-04
34	37	1.5	857 22	T14564	Human milk kappa-case	3.64e-04
35	37	1.5	857 8	Q46850	Recombinant human Kap	3.64e-04
36	37	1.5	906 13	Q80222	Human NDF-beta3 clone	3.64e-04
37	37	1.5	1338 11	Q65607	Rabbit zona pellucida	3.64e-04
38	37	1.5	1348 12	Q78278	T. nlyeum GAPDH.	3.64e-04
39	36	1.5	1502 39	V10099	Human MAP kinase p38-	9.40e-04
40	36	1.5	1737 17	Q99007	Clemokine superfamily	9.40e-04
41	37	1.5	2611 31	T67231	DNAK accessory molecu	3.64e-04
42	36	1.5	2884 11	Q84613	Rat ATR2 receptor clon	9.40e-04
43	37	1.5	3581 10	Q44391	Sequence of murine OS	3.64e-04
44	36	1.5	4091 1	Q04035	Stem cell leukaemia (9.40e-04
45	36	1.5	4199 1	Q05330	Stem cell leukaemia (9.40e-04

ALIGNMENTS

RESULT 1
 ID T25991 standard; cDNA to mRNA; 116 BP.
 AC T25991;
 DE 28-OCT-1996 (first entry)
 DE Human gene signature H00508226.
 KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
 KW human; cloning; mapping; non-biased library; diagnosis; detection;
 KW cell typing; abnormal cell function; ss.
 OS Homo sapiens.
 PN M0514772-AL.
 PD 01-JUN-1995.
 PE 11-NOV-1994; J01916.
 PR 12-NOV-1993; JP-355504.
 PA (MANS/) MATSUBARA K.
 PA (OXDB/) OKUBO K.
 PI Matsubara K, Okubo K;
 DR MPI; 95-206931/27.
 PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
 PT for diagnosis of abnormal cell function, by preparing cDNA that
 PT reflects relative abundance of corresp. mRNA in specific human
 PT tissues
 PS Claim 1: Page 19/6; 2245bp; Japanese.
 CC A single-stranded DNA (or its complementary strand or the corresp.
 CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
 CC given in T19001-T76837 and which is able to hybridise to part of
 CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
 CC sequences were obtained from 3'-directed cDNA libraries prepared
 CC from various human tissues; synthesis of cDNA was initiated from the
 CC 3'-end of mRNA by using poly(I) as the sole primer. Since the 3'-
 CC untranslated sequence is unique to a particular mRNA species, almost
 CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
 CC is constructed so as to reflect accurately the relative abundance of
 CC different mRNAs in the particular tissue from which it was derived.
 CC The appearance frequency of a given GS in a cDNA library can be
 CC determined (seep. using primers and probes derived from the GS
 CC sequences) as a means of diagnosing abnormal cell function or for
 CC recognising different cell types.

```

DR      MPI: 88-279927/40.
PT      Introducing random point mutations into nucleic acids -
PS      by prep of single stranded template, annealing a primer, elongation,
PS      misincorporation, completion of molecules and screening.
CC      Disclosure; P; English.
CC      Random point mutations were introduced into the alpha fragment of
CC      E. coli beta-galactosidase. The wild type sequence was obtained as a
CC      single stranded template and an oligonucleotide was hybridised to
CC      it to generate a popn of DNA molecules which terminate at all
CC      possible nucleotide positions within a specified region. The
CC      variable 3' ends generated in this way are used as primers for
CC      reverse transcriptase. Nucleotides are misincorporated by the
CC      reverse transcriptase and the molecules are completed to forms that can be
CC      amplified and then expressed in a suitable host-vector system.
CC      The sequence covers all 176 diff base substitutions, most of which
CC      occurred singularly in any given mutant.
CC      See also P80575.
SQ      Sequence 204 BP; 21 A; 47 C; 17 G; 11 T; 108 Others;

Query Match 1.98; Score 45; DB 1; Length 204;
Best Local Similarity 20.28; Pred. No. 1.44e-07;
Matches 34; Conservative 70; Mismatches 62; Indels 2; Gaps 2

Db      22 agcyrcaycagcbgcagcygcacbcyrrsgyggcccggygycgcagcycaaycdchv 81
Cp      1426 AGCGACAGTGGCANACAGGGAGACGACAGTGGCCCTGGCCGCCCGAGGGGTGCG-CCA 1368
Db      82 gccgymrtctthyyrrmbhvyrdynrsdaaayccyrrsvlydcynachdhhyvbbb 141
Cp      1367 GAAGCGGGGCTCGACGACACAGTCTGCTGTCGTCATGAGGGC-CACGTCACATGCTCTCT 1309
Db      142 yvvhnhnncncbnhcnvbnhnmhrrnyayrrhdirrdvhcve 189
Cp      1308 GGGGATCTGCCCGAAGATCTCAGATGAGGACGATAGACGACCGCGC 1261

RESULT 4
ID      Q51746 standard; cDNA; 91 BP.
AC      Q51746;
DT      31-MAY-1994 (first entry)
DE      Oligonucleotide probe MK14-A
KM      Oligonucleotide; DNA probe; mycobacteria; disease diagnosis;
KW      88.
OS      Synthetic.
OS      Synthesis.
PN      EP-571911-A.
PD      01-DEC-1993.
PF      24-MAY-1993; 108325
PR      26-MAY-1992; US-889651.
PA      (BECT ) BECTON DICKINSON CO.
PI      Shank DD, Spears PA)
PI      WPI; 93-378844/48.
PT      New oligo:nucleotide probes specific for Mycobacteria - used for
PT      detection and amplification of Mycobacteria nucleic acid in
PT      samples
PS      Claim 3; Page 14; 23pp; English.
CC      Oligonucleotide probe MK14-A consists of nucleotides 5-95 of MK14
CC      (Q51735). It hybridised to all spp. of mycobacteria tested, but
CC      cross reacted to a few non-mycobacterial spp. The probe may
CC      be useful as an initial screen for mycobacterial infection.
CC      See also Q51735-45 and Q51747-59.
SQ      Sequence 91 BP; 5 A; 17 C; 15 G; 4 T;

Query Match 1.88; Score 43; DB 9; Length 91;
Best Local Similarity 2.08; Pred. No. 1.06e-06;
Matches 1; Conservative 46; Mismatches 4; Indels 0; Gaps 0;

Db      10 gsrhshayvvhvbnhbvbnhvvhvvvvhvvhvvhvvhvhyvsv 60
Cp      1650 GGGGACCCCGAGCCCGCCAGCGCTCTCCCGGCGCCTGAGCATTTCCGGG 1700

DB      NBI164 standard. DNA; 204 BP.

```

DR P-PSDB, R65154.
PT Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PS comprising a binding domain and an effector domain
PS Disclosure, Page 35, 255pp, English.
CC Q70466 is a generic DNA sequence used to generate random TSAR (Totally
CC Synthetic Affinity Reagents) peptides. This generic formula can also be
CC represented as follows: X(NNB)11(TGC)(NNB)62(NNB)7(TGC)(NNA)10Y. X
CC and Y are flanking restriction sites (X is not the same as Y) that are
CC not specified further. Other generic sequences are shown in Q70466-68.
CC Other specific peptides generated by these generic sequences are shown in
CC R65151-54. TSARs are conformational heterofunctional proteins or peptides,
CC comprising at least two functional regions - a binding domain with
CC affinity for a ligand and a second effector peptide portion that is
CC chemically or biologically active. They may further comprise a linker
CC peptide between the 2 domains. The oligonucleotides are also designed so
CC that the expressed peptide contains 2 or 4 cysteine residues positioned
CC in, or flanking, the unpredicted or variant residues. These residues
CC confer some degree of conformational rigidity to the peptides. The TSARs
CC or compars, comprising a TSAR binding domain can be used in vivo to
CC deliver a chemically or biologically active moiety, eg. metal ion,
CC radiolabeled peptide, toxin or enzyme, to the specific target or on the
CC cell. They can also replace the function of macromolecules, eg.
CC monoclonal or polyclonal antibodies and therefore circumvent the need
CC for complex methods of hybridoma formation or in vivo antibody
CC production. The TSARs are easily characterised and have designed activity
CC allowing direct and rapid detection in a screening process.
SQ Sequence 114 BP; 0 A; 2 C; 2 G; 2 T;

Query Match 1.7%; Score 40; DB 12; Length 114;
Best Local Similarity 5.4%; Pred. No. 2,03e-05;
Matches 6; Conservative 34; Mismatches 72; Indels 0; Gaps 0;

Dd 3 bmbnbmbnbmbnbmbnbmbnbmbnbmbnbmbnbmbnbmbnnnnnnnnnn 62
QY 225 TGAGGGGAGTGGGCGTTGACCTCTTTGCGAGCGTCATCGACAGATGAGTGAGGGCTGC 284
Dd 63 bmbnbmbnbmbnbmbnbmbnbmbnbmbnbmbnbmbnbmbnbmbnbmbnb 114
QY 285 CCGCGCTCTCTCTCGTCGTCGCCGCCGCTCGGGGGATGCGACATTGCTCTCC 336

RESULT 7
ID T76363 standard; DNA; 172 BP.
AC T76363;
DT 15-SEP-1997 (first entry)
DE Human Interleukin 8 antisense oligonucleotide.
KW Asthma; airway epithelium; adenosine free; cystic fibrosis;
KW chronic obstructive pulmonary disease; bronchitis; ss.
OS Synthetic.
PN H09640162-A1.
PD 19-DEC-1996.
PF 06-JUN-1996; U09306.
PR 07-JUN-1995; US-474497.
PA (UYEC) UNIV EAST CAROLINA.
PI Metzger WJ NYCE JN/
DR MPI; 97-051871/05.

PT Treatment of airway diseases such as asthma - by topically applying
PT adenosine-free antisense oligo:nucleotide to airway epithelium of
PT subject
PS claim 5: Page 36; 71pp; English.
CC A method for treating airway disease in a subject has been produced,
CC which involves the topical administration of an essentially adenosine
CC free antisense oligonucleotide (ON) to the airway epithelium of the
CC subject. The present sequence is an antisense oligonucleotide specific
CC for the human interleukin 8, targeted at the initiation codon. The
CC method can be used to treat airway diseases such as cystic fibrosis,
CC asthma, chronic obstructive pulmonary disease, bronchitis and other
CC airway diseases characterised by an inflammatory response. By
CC administering adenosine from the antisense ON, its liberation upon
CC adenosine degradation is prevented, thereby preventing adenosine-
CC induced bronchoconstriction in patients with hyper-reactive airways.
CC Sequence 172 bp. 1 C; 1 G; 1 T; 1 A.

```

KW Alpha; beta; neu differentiation factor; NDF; human; rat; p185-neu;
KM tyrosine phosphorylation; differentiation; phenotype; proliferation;
KM wound; tumour; epithelial tissue; breast; stomach;
KM gastrointestinal disease; Barrett's oesophagus;
KM (non-)cystic kidney disease; inflammatory bowel disease; ds.
OS Rattus rattus.
FH Key Location/Qualifiers
FT cds 17..712 /*tag- a
FT /product= Rat NDF
PD MO9428133-A.
PD 08-DEC-1994.
PE 23-MAY-1994; U05769.
PR 21-MAY-1993; US-066384.
PA (AMGE-) AMGEN INC.
PI Hu S, Koshl RA, Liu N, Pierce GF, Sugarman BJ;
PI Men D,
DR WPI: 95-022805/03.
DR P-PADB; R68570.
PT New recombinant neu differentiation factors and corresp. DNA -
PT are used in the treatment of tumours, dermal wounds, and
PT gastrointestinal, kidney and inflammatory bowel diseases.
PS Disclosure: Page 170-173, 311pp, English.
CC The sequences given in 080233-93 encode: rat neu differentiation
CC factors (NDF's). The peptides encoded by these cDNA clones isolated
CC from rat tissues and cell lines, possess the ability to stimulate
CC human p185-neu tyrosine phosphorylation. These peptides have the
CC ability to induce a differentiated phenotype in certain cell lines and
CC can stimulate or inhibit proliferation of certain cell lines. The
CC NDF's can be used to treat wounds, tumours derived from epithelial
CC tissue of the breast, stomach etc., gastrointestinal disease, Barrett's
CC oesophagus, (non-)cystic kidney disease or inflammatory bowel disease.
CC These cDNA sequences were used in the isolation of related sequences
CC from human cDNA libraries which encode human NDF's.
SQ Sequence 1098 BP; 343 A; 265 C; 254 G; 236 T;

Query Match 1.7%; Score 41; DB 13; Length 1098;
Best Local Similarity 81.5%; Pred. No. 7,63e-06;
Matches 53; Conservative 0; Mismatches 12; Indels 0; Gaps 0

Db 1034 tggtaataaaatcatcccaacaagttacaaaataaaaaaaaaaaaaaaaggcg 1093
||| ||||||| ||||| ||| ||||||| ||||| ||||| ||||| ||||| |||||
QY 2326 TGCGGAAATPAAAGCATTTCAGNGCAAAAAAAAAAAAAAAAAAAAAAAGCGCG 2385
|||||
DB 1094 ccgcg 1098
|||||
QY 2386 GCCGC 2390

RESULT 10
ID 080218 standard; DNA; 1651 BP.
AC 080218:
DE 31-AUG-1995 (first entry)
Human NDF-alpha2b clone 17 DNA.
KW Alpha; beta; neu differentiation factor; NDF; human; rat; p185-neu;
KM tyrosine phosphorylation; differentiation; phenotype; proliferation;
KM wound; tumour; epithelial tissue; breast; stomach;
KM gastrointestinal disease; Barrett's oesophagus;
KM (non-)cystic kidney disease; inflammatory bowel disease; ds.
OS Homo sapiens.
FH Key Location/Qualifiers
FT cds 20..820 /*tag- a
FT /product= Human NDF-alpha2b
PD MO9428133-A.
PD 08-DEC-1994.
PE 23-MAY-1994; U05769.
PR 21-MAY-1993; US-066384.
PA (AMGE-) AMGEN INC.
PI Hu S, Koski RA, Liu N, Pierce GF, Sugarman BJ;
PI Men D,
DR WPI: 95-022805/03.
DR P-PADB; R68563.

```

PT New recombinant neu differentiation factors and corresp. DNA -
 PT are used in the treatment of tumours, dermal wounds, and
 CC gastrointestinal, kidney and inflammatory bowel diseases.
 PS Claim 25: Page 139-141; 341pp; English.
 CC The sequences given in 080215-22 encode human neu differentiation
 CC factors (NDF's). The peptides encoded by these cDNA clones isolated
 CC from human tissues and cell lines, possess the ability to stimulate
 CC human p185-neu tyrosine phosphorylation. These peptides have the
 CC ability to induce a differentiated phenotype in certain cell lines and
 CC can stimulate or inhibit proliferation of certain cell lines. The
 CC NDF's can be used to treat wounds, tumours derived from epithelial
 CC tissue of the breast, stomach etc., gastrointestinal disease, Barrett's
 CC oesophagus, (non-)cystic kidney disease or inflammatory bowel disease.
 CC These cDNA sequences may be used in a recombinant plasmid for the
 CC recombinant production of the NDF's in a pro- or eukaryotic host cell.
 SO Sequence 1651 BP; 510 A; 460 C; 352 G; 349 T;

Query Match 1.7%; Score 40; DB 13; Length 1651;
 Best Local Similarity 100.0%; Pred. No. 2.03e-05;
 Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1612 aaaaaaaaaaaaaaaaaaaaaaaaaagggcgccgc 1651
 ||||||||||||||||||||||||||||||||||||||||
 OY 2351 AAAAAAAAAAAAAAAAAAAAAAAAAAGGCGCGCCG 2390

RESULT 11
 ID 139792 standard; DNA; 1734 BP.
 AC T39792;
 DT 19-FEB-1998 (first entry)
 DE Human SH3P9 gene.
 KW Src-homology region 3 domain; human; mouse; SH3 domain; cell growth;
 KW cellular signalling element; cellular structural element; malignancy;
 KW protein identification; functional domain; protein screening;
 KW cellular signal transduction process; ss.
 OS Homo sapiens.
 FH Key
 FT CDS Location/Qualifiers
 FT 1..1215
 FT /tag= a
 FT /product= human SH3P9
 PN WO9631625-A1.
 PD 10-OCT-1996.
 PF 04-APR-1996; U04454.
 PR 03-APR-1996; US-630915.
 PR 07-APR-1995; US-417872.
 PA (CYTO-) CYTOGEN CORP.
 PA (UYNC-) UNIV NORTH CAROLINA.
 PI Fowlkes DM, Hoffman N, Kay BK, McConnell SJ, Sparks AB;
 DR WPI; 96-465045/46.
 DR P-PSDB; W03392.
 PT Identifying polypeptide(s) having specific functional domain (esp.
 PT SH3 domain) - comprises detecting selective binding to recognition
 PT unit, regardless of sequence homology
 PS Claim 76: Fig 36; 17pp; English.
 CC T39786-T39803 represent novel human and mouse genes encoding Src-homology
 CC region 3 (SH3) domain containing proteins that can be used in the method
 CC of the invention. SH3 domain containing proteins play a role in
 CC signalling and structural elements of cells. The method of the invention
 CC is for identifying polypeptides containing functional domains of interest
 CC (especially SH3 domains). The method comprises contacting a multivalent
 CC recognition unit (RU) complex with a number of peptides and identifying
 CC polypeptides having a selective binding affinity for the RU complex. The
 CC method is based on functional similarities and does not rely on sequence
 CC similarities. Prior methods only gave limited success for identifying
 CC proteins which contain an SH3 domain due to the minimal sequence homology
 CC among known SH3 proteins. It has been found that small peptide RUs in
 CC multivalent form have reduced specificity for a given functional domain
 CC compared to monomer RUs. Multivalent RU complexes are particularly suited
 CC to screening for polypeptides containing functional domains that are
 CC similar to, but not identical in sequence to, the original target
 CC functional domain. The new method enables proteins having a common
 CC function to be identified. Identification of novel SH3 proteins will be
 CC useful for a better understanding of cell growth, malignancy, signal

CC transduction processes, etc. New candidate drugs can be identified, and
 CC their specificities (e.g. pharmacological activities) can be assessed
 CC using the method of the invention.
 SO Sequence 1734 BP; 418 A; 494 C; 526 G; 296 T;

Query Match 1.7%; Score 40; DB 35; Length 1734;
 Best Local Similarity 100.0%; Pred. No. 2.03e-05;
 Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1695 aaaaaaaaaaaaaaaaaaaaaaaaaagggcgccgc 1734
 ||||||||||||||||||||||||||||||||||||||||
 OY 2351 AAAAAAAAAAAAAAAAAAAAAAAAAAGGCGCGCCG 2390

RESULT 12
 ID 080228 standard; DNA; 3344 BP.
 AC 080228;
 DT 31-AUG-1995 (first entry)
 DE Rat NDF clone 22 DNA.
 KW Alpha; beta; neu differentiation factor; NDF; human; rat; p185-neu;
 KW tyrosine phosphorylation; differentiation; phenotype; proliferation;
 KW wound; tumour; epithelial tissue; breast; stomach;
 KW gastrointestinal disease; Barrett's oesophagus;
 KW (non-)cystic kidney disease; inflammatory bowel disease; ds.
 OS Rattus rattus.
 FH Key
 FT CDS Location/Qualifiers
 FT 345..2255
 FT /tag= a
 FT /product= Rat NDF
 PN WO9428133-A.
 PD 08-DEC-1994.
 PF 23-MAY-1994; U05769.
 PR 21-MAY-1993; US-066384.
 PA (AMGE-) AMGEN INC.
 PI Hu S, Koeki RA, Liu N, Pierce GF, Sugarman BJ;
 PI Wen D;
 DR WPI; 95-022805/03.
 DR P-PSDB; R68573.
 PT New recombinant neu differentiation factors and corresp. DNA -
 PT are used in the treatment of tumours, dermal wounds, and
 PT gastrointestinal, kidney and inflammatory bowel diseases.
 PS Disclosures: Page 184-188; 341pp; English.
 CC The sequences given in 080223-33 encode rat neu differentiation
 CC factors (NDF's). The peptides encoded by these cDNA clones isolated
 CC from rat tissues and cell lines, possess the ability to stimulate
 CC human p185-neu tyrosine phosphorylation. These peptides have the
 CC ability to induce a differentiated phenotype in certain cell lines and
 CC can stimulate or inhibit proliferation of certain cell lines. The
 CC NDF's can be used to treat wounds, tumours derived from epithelial
 CC tissue of the breast, stomach etc., gastrointestinal disease, Barrett's
 CC oesophagus, (non-)cystic kidney disease or inflammatory bowel disease.
 CC These cDNA sequences were used in the isolation of related sequences
 CC from human cDNA libraries which encode human NDF's.
 SO Sequence 3344 BP; 962 A; 871 C; 749 G; 762 T;

Query Match 1.7%; Score 40; DB 13; Length 3344;
 Best Local Similarity 100.0%; Pred. No. 2.03e-05;
 Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 3305 aaaaaaaaaaaaaaaaaaaaaaaaaagggcgccgc 3344
 ||||||||||||||||||||||||||||||||||||||||
 OY 2351 AAAAAAAAAAAAAAAAAAAAAAAAAAGGCGCGCCG 2390

RESULT 13
 ID V02308 standard; cDNA; 3451 BP.
 AC V02308;
 DT 06-MAY-1998 (first entry)
 DE Cell membrane proton-ATPase encoding cDNA.
 KW Cell membrane proton-ATPase; Cyanidium caldarium; acid resistance;
 KW transgenic plant; ds.
 OS Cyanidium caldarium.
 FH Key
 FT Location/Qualifiers

```

FT FT CDS 340..3307
FT FT /*tag= a
FT FT /product= "cell membrane proton-ATPase"
PD PD J09252786-A.
PD PD 30-SEP-1997.
PF PF 25-MAR-1996; 124124.
PF PF 25-MAR-1996; JP-124124.
PA (ORF1) ORIENTAL YEAST CO LTD.
DR WP1: 97-530157/49.
DR P-PSDB: W31363.
PT Proton-ATPase and cDNA coding sequence - useful for producing
PS acid-resistant transgenic plants
PS Claim 1; Page 11-13; 16pp; Japanese.
CC The present sequence encodes a cell membrane proton-ATPase.
CC The gene is useful for producing acid-resistant transformed
CC plants. It also allows large scale production of the enzyme.
SQ Sequence 3451 BP; 818 A; 860 C; 960 G; 813 T;

Query Match 1.7% Score 40; DB 38; Length 3451;
Best Local Similarity 100.0%; Pred. No. 2,03e-05;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0

Db 3412 aaaaaaaaaaaaaaaaaaaagggcgccgcg 3451
    |||||||
Oy 2351 AAAAAAAAAAAAAAAAAAAAGGCGCCGCC 2390

RESULT 14
ID 070466 standard; DNA; 114 BP.
AC 070466;
DT 05-APR-1995 (first entry)
DE Generic DNA sequence to generate a random TSAR-9 petide library.
DE TSAR: totally synthetic affinity reagent; synthetic; binding domain;
KW effector domain; concatenated heterofunctional protein; linker;
OS direct; rapid; detection; screening; treatment; generic; ss.
FH Key
FT misc_feature
FT 55..60
FT Location/Qualifiers
FT /*tag= a
FT /note= "this sequence represents 'Z' 1 2 can be a
FT sequence of 6, 9 or 12 nucleotides (see
FT comments)".
PD WO9418318-A.
PD 18-AUG-1994.
PF 01-FEB-1994; U00977.
PF 01-FEB-1993; US-013416.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PR (UYNC-) UNIV NORTH CAROLINA.
PA Fowlkes DM, Kay BK.
PA WP1: 94-279739/34.
DR P-PSDB: R05152.
PT Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins
PT comprising a binding domain and an effector domain
PS Disclosure: Page 35; 255pp; English.
O70466 is a generic DNA sequence used to generate random TSAR (Totally
CC Synthetic Affinity Reagents) peptides. This generic formula can also be
CC represented as follows: X(NNB)1(TGC)(NNB)10(TGC)2(NNB)42(NNB)8(TGC)(NNB)
CC -97. X and Y are flanking restriction sites (X is not the same as Y)
CC 070466-66. Other specific peptides generated by these generic sequences
CC are shown in R65151-54. TSARs are concatenated heterofunctional proteins
CC or peptides, comprising at least two functional regions - a binding
CC domain with affinity for a ligand and a second effector peptide portion
CC that is chemically or biologically active. They may further comprise a
CC linker peptide between the 2 domains. The oligonucleotides are also
CC designed so that the expressed peptide contains 2 or 4 cysteine residues
CC positioned in, or flanking, the unpredicted or variant residues. These
CC residues confer some degree of conformational rigidity to the peptides.
CC The TSARs or compns, comprising a TSAR binding domain can be used in
CC vivo to deliver a chemically or biologically active moiety, eg metal
CC ion, radioisotope, peptide, toxin or enzyme, to the specific target or

```

```
CC      on the cell. They can also replace the function of macromolecules, eg.  
CC      monoclonal or polyclonal antibodies and therefore circumvent the need for  
CC      complex methods of hybridoma formation or in vivo antibody production.  
CC      The TSARs are easily characterised and have designed actively allowing  
CC      direct and rapid detection in a screening process.
```

SQ Sequence 114 BP; 0 A; 4 C; 4 G; 4 T;

```
Query Match
Best Local Similarity 1.6%; Score 39; DB 12; Length 114;
Matches 9; Conservative 30; Mismatches 67; Indels 0; Gaps 0;
```

Dn dbnbnbnbnbnbnbnbnbnbnbnbnbtgttcnnbnbnbnbnbnbnbnbnbnbnbn

Oy 742 CTGCTGTTCCTCGTGGCCTTTCTCCGCCGAGTGCTGCATTCACGCCGCCGCCCCC 801

Dd 69 dnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbn 114

Oy 802 GGCTCCCTGAGAGTGTCTATTCTTGGGCTTCACGCTGCTGT 847

RESULT 15

ID T76405 standard; DNA; 178 BP.

DC T76405;

DT 15-SEP-1997 (first entry)

KD Human endothelin-1 antisense oligonucleotide.

NW Acthma; airway epithelium; adenoseis free; cystic fibrosis;
KW chronic obstructive pulmonary disease; bronchitis; ss.
OS Synthetic.
PN W09640162-AI.
PD 19-DEC-1996.
PF 06-JUN-1996; D09306.
PR 07-JUN-1995; US-474497.
PA (UYEC-) UNIV EAST CAROLINA.
PI Metzger MJ NYCE JW;
WPJ: 97-051871/05.
PT Treatment of alrwy diseases such as asthma - by topically applying
PT adenosine-free antisense oligo:nucleotide to alrwy epithelium of
PT subject

PS Claim 5: Page 38; 71pp; English.

CC A method for treating airway disease in a subject has been produced,
CC which involves the topical administration of an essentially adenosine
CC free antisense oligonucleotide (ON) to the airay epithelium of the
CC CC subject. The present sequence is an antisense oligonucleotide specific
CC for the human endotheln-1 targeted at the initiation codon. The
CC method can be used to treat airway diseases such as cystic fibrosis,
CC astma, chronic obstructive pulmonary disease, bronchitis and other
CC alirwy diseases characterised by an inflammatory response. By
CC eliminating adenosine from the antisense ON, its liberation upon
CC antisenase degradation is prevented, thereby preventing adenosine-
CC induced bronchoconstriction in patients with hyper-reactive airways.

SQ Sequence 178 BP; 0 A; 52 C; 46 G; 32 T;

```
Query Match
Best Local Similarity 1.6%; Score 39; DB 32; Length 178;
Matches 24; Conservative 31; Mismatches 16; Indels 0; Gaps 0;
```

Dd 105 cbbgcacbbbcbgbgdbggbbbbtcttgagcabbttbcctctqbgbbbbbbsggat 164

Oy 1919 CCCGCTCCCTGGGGGTGGGTGCCCAAGACCCTGAAGCCGCTCTGCTTGCCTGCCCCAAGTGGG 1978

Dd 165 cbdbbbcctcc 175

Oy 1979 CGCGCACCCCC 1989

Search completed: Sun Feb 14 21:36:35 1999

Job time : 361 secs.

 RELEASE 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

MPearch_n n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Sun Feb 14 19:34:43 1999; Maspar time 3738.48 Seconds
 1512.673 Million cell updates/sec

Tabular output not generated.

Title: >US-08-946-869-9
 Description: (1-2393) from US08946869.seq
 Perfect Score: 2393
 N.A. Sequence: 1 AAGCCACCTCCGCGCTCTC.....AAAAAAGCGCGCGCGGT 2393
 Comp: TTCGGGTGAGCGCGCGAGAG.....TTTTCCTCCGCGCGCGCA

Scoring table: TABLE default
 Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 602357 seqs, 1181590623 bases x 2

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: emb157
 1:em_ba 2:em_fun 3:em_htg 4:em_hum1 5:em_hum2 6:em_in
 7:em_com 8:em_or 9:em_ov 10:em_pat 11:em_ph 12:em_pl
 13:em_ro 14:em_un 15:em_vl
 Database: genbank110
 16:gb_pat 17:gb_pat2 18:gb_htg 19:gb_in 20:gb_com 21:gb_ov
 22:gb_pat 23:gb_ph 24:gb_pl 25:gb_pl2 26:gb_plt
 27:gb_pat2 28:gb_pat3 29:gb_ro 30:gb_st 31:gb_sts 32:gb_sy
 33:gb_un 34:gb_vl

Statistics: Mean 12.262; Variance 12.056; scale 1.017

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	258	10.8	271 31	G22461	human STS WI-30695.	6,03e-89
2	118	4.9	6220 28	AB001535	Human sapiens mRNA, com	5,59e-31
3	86	3.6	78064 18	HSAC000376	*** SEQUENCING IN PROG	2,11e-18
4	86	3.6	155074 28	AC003693	Human Chromosome 11p15	2,11e-18
5	78	3.3	7218 22	166494	Sequence 14 from paten	2,39e-15
6	54	2.3	7218 22	166494	Sequence 14 from paten	1,51e-06
7	53	2.3	78064 18	HSAC000376	*** SEQUENCING IN PROG	3,39e-06
8	45	1.9	148869 18	AC006037	*** SEQUENCING IN PROG	1,86e-03
9	45	1.9	193381 18	AC005158	*** SEQUENCING IN PROG	1,86e-03
10	40	1.7	1016 25	AF024520	Arabidopsis thaliana M	8,08e-02
11	40	1.7	2285 29	RN30290	Mus musculus gamma-ant	8,08e-02
12	40	1.7	2325 29	RN30290	Rattus norvegicus gale	8,08e-02
13	40	1.7	2865 27	HS181867	Homo sapiens (subclone	8,08e-02

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
14	40	1.7	3451 22	E13998	Cyanidium caldarium mR	8,08e-02
15	40	1.7	3519 27	HS181869	Homo sapiens (subclone	8,08e-02
16	40	1.7	6922 29	MA017793	Mus musculus mRNA for	8,08e-02
17	40	1.7	74371 27	AC005369	Human sapiens chromosom	3,85e-02
18	40	1.7	75270 28	AF023268	Homo sapiens c1k2 kina	3,85e-02
19	41	1.7	75270 28	HS377516	Human DNA sequence ***	3,50e-01
20	38	1.6	965 22	AR024429	Sequence 22 from paten	1,69e-01
21	39	1.6	965 22	AR024429	Sequence 22 from paten	1,69e-01
22	38	1.6	1024 20	BF079414	Bos taurus common sal	3,50e-01
23	38	1.6	1285 24	ADP619	A. deliliosa p91p mRNA	3,50e-01
24	38	1.6	1425 22	AR005195	Sequence 1 from patent	3,50e-01
25	38	1.6	1425 22	181228	Sequence 1 from patent	3,50e-01
26	38	1.6	1425 22	181228	Sequence 1 from patent	3,50e-01
27	38	1.6	1451 27	AF013988	Homo sapiens serine pr	1,59e-01
28	39	1.6	1659 20	AF024943	Sus scrofa growth horm	3,50e-01
29	38	1.6	1799 24	AF022082	Arabidopsis thaliana 8	3,50e-01
30	38	1.6	2399 22	A07647	Synthetic gene for pla	3,50e-01
31	38	1.6	2573 19	DM049724	Drosophila melanogaste	3,50e-01
32	38	1.6	2775 21	PAF0AT	Pseudopneumocystis ame	3,50e-01
33	38	1.6	2862 29	126126	angiotensin II type 2	3,50e-01
34	38	1.6	2862 29	126126	Sequence 1 from patent	3,50e-01
35	38	1.6	3433 27	HS338110	Human DNA sequence fro	1,69e-01
36	39	1.6	35730 27	HSN7467	Human DNA sequence fro	3,50e-01
37	38	1.6	35730 27	AC005369	Human DNA sequence ***	1,69e-01
38	39	1.6	74371 27	AC005369	Human DNA sequence ***	1,69e-01
39	39	1.6	151840 18	HS919B11	Human sapiens chromosom	1,69e-01
40	39	1.6	169144 18	AC005849	*** SEQUENCING IN PROG	1,69e-01
41	39	1.6	169144 18	AC005849	Human sapiens chromosom	1,69e-01
42	41	1.6	183827 28	AC005899	Human sapiens sequence ***	7,20e-01
43	39	1.6	212175 18	HS75M14	Human DNA sequence ***	7,20e-01
44	37	1.5	130027 27	AC004982	Homo sapiens P4 clone	7,20e-01
45	37	1.5	189666 18	AC004052	*** SEQUENCING IN PROG	7,20e-01
46	37	1.5	220000 18	AC004480	*** SEQUENCING IN PROG	7,20e-01

ALIGNMENTS

RESULT 1 G22461 271 bp DNA STS 31-MAY-1996
 LOCUS human STS WI-30695.
 DEFINITION G22461
 ACCESSION G1342787
 KEYWORDS STS sequence; primer; sequence tagged site.
 SOURCE human STS derived from sequences in dbEST and the UniGene
 collection.

ORGANISM Homo sapiens
 Eukaryote; mitochondrial eukaryotes; Metazoa; Chordata;
 Vertebrate; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 REFERENCE Hudson, R.
 AUTHORS Whitehead Institute/MIT Center for Genome Research; Physically
 TITLE Mapped STS
 JOURNAL Unpublished (1995)
 COMMENT

Contact: Thomas Hudson
 Whitehead Institute/MIT Center for Genome Research
 Whitehead Institute for Biomedical Research
 9 Cambridge Center, Cambridge MA 02142 USA
 Tel: 617 252 1900
 Fax: 617 252 1902
 Email: thudson@genome.wi.mit.edu

Primer A: CCCAGTGTGAGAAATCTGT
 Primer B: CTGGAGGATCAAGGCTT
 STS size: 112
 PCR Profile:

Denaturation: 96 degrees C
 Annealing: 56 degrees C
 Polymerization: 72 degrees C
 PCR Cycles: 35
 Thermal Cycler:
 Protocol:
 Template: 10 ng

US-08-946-869-9.rge

Page 2

```

FEATURES
source
Japan (E-mail: jshimizu@med.koto.ac.jp, Tel: 03-3351-2370,
Fax: 03-3351-2370)
Location/Qualifiers
1. 6220
/organism="Homo sapiens"
/note="7 Caucasians"
/db_xref="taxon:9606"
/chromosome="21"
/dev_stage="20-26weeks fetus"
/map="21q22.3"
/clone_lib="human fetal brain 5'-STRECH plus cDNA library
(CLOMTECH) and Human Brain, caudate nucleus 5'-STRECH cDNA
library (CLOMTECH)"
/clone_type="brain"
446..4957
/gene="KNP3"
446..4957
/gene="KNP3"
/note="similar to C.elegans hypothetical protein
CEP01H8.1, CE05C12.3, CE54D1.5. similar to tnp and
tvp-like proteins"
/codon_start=1
/db_xref="pid:d0103683"
/db_xref="pid:g3928756"
/translation="MEPSALRAGSEDEGEFGLPRVTDGAVNLSRNSLSFSG
RACPGFNDKQESLSWIPENIKKKEVYFESSKLSDAKRVCCGGYTHECHLEA
TRPPTQMDPKRKHVOEMPTDARGLVPGISQNKYKRVASQDTPSSVTHLMLA
HNGVDNLLISVTGAGAKNENPKLSIRGLYKVAQTGAVIITGSHGTVMKV
GAVNDPFLSSSVKREGELITIGVATWGVHNRRLIHPGTFPAEYILDDGQNLTC
LDSNSHPTIYNDGTHGQIVGVEITRLEPRLIEPTEKRGVAKIKIVCYLEGP
GLTHIDNATNGTQPCVVEGSGRAVDAIVANLPSVDITSLLOOKSVFQEMFE
TRESRIYEMTKIDIVRRQLITVPEEGDGOODVAILOHLAASNSDQFHE
MNDHDLVAVNRVDIARSELIPDENOMKPSDHPNMAALISNPEFKMLLENGV
QKEVEVTLVLYLENDDPSCLPHSKLOAVDEPRPACAPAAPILOMHRVAVLA
ELIGDQPIYPRPRHNDRLILPVPHVNLVGVSAKQAILHNRBRVDLPRGAVY
LIIAVTQNRRELDEIIWAOSODCIAAALSKILKEEEDTSPSEMLALAEY
EHRALIVTECTRKDEEARKOKLRVSEANGTKTCLQIALAENDKPVSHGGLAVLT
RYVMQGLSVNDGLKRAVILCMALFPLITGLISPERLDOVGPILAAARAFIAPV
PHNLSTISPAFLCEAVLYAMDQPVPSKCECAIYMLFSEVKEEROLPYDEGCL
NMKALYPSDFPMNLQVAILIPAGLITCLITRATYIPGVVILSDPILFCLRHIF
TISTGQIKIITVRKAMDFFELFLAAVVSQVAKQAILHNRBRVDLPRGAVY
HSYLTIPQIIGTIDGVNTPENRSPGNDPYPKPCESDQAQORAPAPPEVTLIC
LYLFTNLILNLINAEVNTFOVDERTQYKFORHDIIEYHNPAPAPPEVTLIC
HLOETIKRVVLTTPAKRHKOKNKLSENEAALISREYIKENTLONRPOQOREO
KIEDIBSNKVDAMDLDLDPLKRSGEQGLALBEDVATATLMLIVRTLASGFS
SEADVILASOKAAEEDBAPGKRKTKEEGLALBEDVATATLMLIVRTLASGFS
PWMEETIYDPPYTPARKNMAAMDPIQDTELESTIOYNVVGLDRLSPFNPQVY
AGLEINMKGRLGKRGRLSCFQPNHILPMTVRMRNEGALCRSIRKMLLEVLYK
LPSEHNAIPGSRPEPEMLPRILKRLIRQEHSPENLTKGMPEYKGYKMDLPRTD
NAMLEIYAVVHPQDNDVELNRLNSLNACDSGASIRHGVYDRLPIYANHKTLQK
AAAEFGAY"
BASE COUNT 1248 a 1938 c 1882 g 1152 t
ORIGIN
Query Match 4.9% Score 118, DB 28; Length 6220;
Beat Local Similarity 60.6%; Pred. No. 5.59e-11;
Matches 423; Conservative 0; Mismatches 269; Indels 6; Gaps 4;
Db 3032 TACTTCAGTCACTTCGAAATPAAGCTGAGAGTGGGGCAATCTTCTTCGTGGCAGG 3091
QY 945 TACCTCGCGCAGACTGGAACAGTGGAGACTGTGGCTTCACCTGCTCTCCCGGC 1004
3092 CTGACCTGGCAGGCTCATCCGGCGAGAGCTGTATCCCGGCGCGATCTCTCTGTGCAC 3151
QY 1005 GTGGGCTGCGCGGTGAGACCCCGGGGTGTGTGACACACCTGGCGGCACTGTCTGCACTGAC 1064
Db 3152 TTCACTCTGTCTGCGCTCCGGGCTCAGCAGCATTTTACCATGATPAAGACGTGGCGCC 3211
QY 1065 TTCAATGGTTTTCAGACGCTGGCGGCTTCACACTTCACCGGTCAACAAACAGCTGGCGCC 1124
Db 3212 AAGATCATCATTTGTGAAAGCGATGATGAAAGACGCTCTCTCTCTCTCTGCTGCTGCT 3271

```


MUSE RELEASE

(TM)

Release 3.1a John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MParch_n n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Sun Feb 14 21:30:34 1999; Mspar time 353.20 Seconds

Tabular output not generated. 921.792 Million cell updates/sec

Title: >US-08-946-869-9
(1-2393) from US08946869.seq

Perfect score: 2393

N.A. Sequence: 1

Comp: 1 AGGCGACCTCGCGCTCTC.....AAAAAGGCGCGCGCGT 2393
TTCGCGTGAGCGCGAGAG.....TTTTTCCCGCGCGCGCA

Scoring table: TABLE default

Gap 6

Mismatch STD: Dbase 0; Query 0

Searched: 188442 seqs, 68026449 bases x 2

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database:

n-geneseq32
1:part1 2:part2 3:part3 4:part4 5:parts 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22 23:part23
24:part24 25:part25 26:part26 27:part27 28:part28
29:part29 30:part30 31:part31 32:part32 33:part33
34:part34 35:part35 36:part36 37:part37 38:part38
39:part39 40:part40

Statistics: Mean 9.947; Variance 7.935; scale 1.253

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	116	4.8	116	22	Human gene signature	1.45e-41
2	48	2.0	91	9	Oligonucleotide probe	6.92e-09
3	45	1.9	204	1	Base substituted E.co	1.44e-07
4	43	1.8	91	9	Oligonucleotide probe	1.06e-06
5	42	1.8	204	1	Base substituted E.co	2.85e-06
6	40	1.7	114	12	Generic DNA sequence	2.03e-05
7	40	1.7	172	33	Human interleukin 8 a	2.03e-05
8	40	1.7	946	13	Human NDF- α phaz2 clone	7.63e-06
9	40	1.7	1098	13	Rat NDF clone 4 DNA	2.03e-05
10	40	1.7	1651	13	Human NDF- α phaz2 clone	2.03e-05
11	40	1.7	1734	35	Human SHP9 gene	2.03e-05
12	40	1.7	3344	13	Rat NDF clone 22 DNA	2.03e-05
13	40	1.7	3451	38	Cell membrane protein	2.03e-05

ID	Accession	Score	Query Match	Length	Description	Pred. No.
14	38	1.6	114	12	Generic DNA sequence	1.40e-04
15	39	1.6	178	32	Human endochelin-1 an	5.35e-05
16	38	1.6	359	38	Human secreted protei	1.40e-04
17	38	1.6	359	39	3' portion of cDNA c1	1.40e-04
18	38	1.6	1435	23	Cytoplasmic anti-prote	1.40e-04
19	39	1.6	2335	13	Human PROMPF- α phaz2	5.35e-05
20	38	1.6	2400	1	Placenta-specific pro	1.40e-04
21	38	1.6	2862	14	Mouse A12 receptor cd	1.40e-04
22	33	1.5	56	9	Sequence of oligo nuc	9.40e-04
23	36	1.5	114	12	Generic DNA sequence	9.40e-04
24	36	1.5	114	12	Generic DNA sequence	9.40e-04
25	36	1.5	114	12	Generic DNA sequence	9.40e-04
26	36	1.5	114	12	Generic DNA sequence	9.40e-04
27	36	1.5	162	32	Human RNASES antisens	9.40e-04
28	36	1.5	201	38	3' fragment of clone	9.40e-04
29	36	1.5	607	38	Human secreted protei	9.40e-04
30	36	1.5	607	39	Human secreted protei	9.40e-04
31	36	1.5	607	39	Human secreted protei	9.40e-04
32	36	1.5	607	39	Human secreted protei	9.40e-04
33	36	1.5	701	32	Mouse PLAZS cDNA (B6a	9.40e-04
34	37	1.5	810	32	Mouse PLAZS cDNA (v11	9.40e-04
35	37	1.5	857	22	Human milk kappa-case	3.64e-04
36	37	1.5	857	8	Human recombinant h	3.64e-04
37	37	1.5	906	13	Human NDF-beta3 clone	3.64e-04
38	37	1.5	1338	11	Rabbit zona pellucida	3.64e-04
39	36	1.5	1348	12	T. nilemus GAD6H	3.64e-04
40	36	1.5	1502	39	Human MAP kinase p38-	9.40e-04
41	37	1.5	1737	17	Chemokine superfamily	9.40e-04
42	37	1.5	2611	31	DNAX accessory molecu	9.40e-04
43	36	1.5	2884	14	Rat A72 receptor clon	9.40e-04
44	37	1.5	3581	10	Sequence of murine OS	9.40e-04
45	36	1.5	4091	1	Stem cell leukaemia (9.40e-04
			4199	1	Stem cell leukaemia (9.40e-04

ALIGNMENTS

RESULT 1
ID T25991 standard; cDNA to mRNA; 116 BP.
AC T25991;
DT 28-OCT-1996 (first entry)
DE Human gene signature HUMG08326.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
OS cell typing; abnormal cell function; ss.
KW Homo sapiens.
PN NC0514772-A1.
PD 01-JUN-1995.
PE 11-NOV-1994; 019136.
PR 12-NOV-1993; JP-355504.
PA (MATS//) MATSUBARA K.
PI (OKUB//) OKUBO K.
DR WPI; 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1, Page 1976; 2245p; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 'GS' sequences
CC given in T19001-726837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (gene signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.

Lahentoivaa P, Knowles J, Kolvula A, Bamford J, Reinikainen T

RESULT 5
ID N81164 standard; DNA; 204 BP.



Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Mpsrch_un n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Sun Feb 14 20:37:53 1999; Mspart time 3113.83 seconds

Tabular output not generated. 1376.223 Million cell updates/sec

Title: >US-08-946-869-9
(1-2393) from US08946869.seq
Description: 2393
Perfect Score: 1 AAGCCACCTGCGCGCTCTC.....AAAAAAGGCGCGCGCGGT 2393
N.A. Sequence: 1 TTCCGCTGACGCGCGAGAG.....TTTTCCTCCCGCGCGCGCA
Comp:

Scoring table: TABLE default
Gap 6

Nmatch STD: Dbase 0; Query 0

Searched: 2275026 seqs, 895388244 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: emb1-est56
1:em-est1 2:em-gs81 3:em-gs82 4:em-gs83
Database: genbank-est109
5:gb-est1 6:gb-est10 7:gb-est11 8:gb-est12 9:gb-est13
10:gb-est14 11:gb-est15 12:gb-est16 13:gb-est17
14:gb-est18 15:gb-est19 16:gb-est20 17:gb-est21
18:gb-est22 19:gb-est23 20:gb-est24 21:gb-est25 22:gb-est26
23:gb-est27 24:gb-est28 25:gb-est29 26:gb-est30 27:gb-est31
28:gb-est32 29:gb-est33 30:gb-est34

Statistics: Mean 12.350; Variance 5.384; scale 2.294

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description	Pred. No.
C 1	470	19.6	495	8	AA551759	nf99c01.s1 NCI_CGAP_Co	0.00e+00
C 2	414	17.3	441	10	AA633315	qh58f08.s1 NCI_CGAP_Co	0.00e+00
C 3	402	16.8	444	11	AA694490	ah33h08.s1 Soares test	0.00e+00
C 4	396	16.5	400	11	AA708532	2163d12.s1 Soares preg	0.00e+00
C 5	377	15.8	381	13	AA878567	0117b06.s1 NCI_CGAP_K1	0.00e+00
C 6	379	15.8	413	19	N31660	yx65b04.s1 Homo sapien	0.00e+00
C 7	350	14.6	410	16	H18835	ym45d10.s1 Homo sapien	7.30e-292
C 8	301	12.6	311	17	AI092908	qa36c03.x1 Soares Nihm	7.78e-245
C 9	297	12.4	309	15	AA932133	om90h09.s1 NCI_CGAP_K1	5.32e-241
C 10	258	10.8	271	16	H18836	ym45d10.s1 Homo sapien	7.99e-204
C 11	257	10.7	294	6	AA026974	ZK01e10.s1 Soares preg	7.12e-203
C 12	255	10.7	299	7	AA454774	Zx77b07.s1 Soares ovar	5.66e-201
C 13	249	10.4	259	10	AA592910	nm01f07.s1 NCI_CGAP_Co	2.80e-195

RESULT LOCUS	DEFINITION	ACCESSION	NID	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1	AA551759	495 bp	MRNA	EST	11-AUG-1997	11-AUG-1997	1	AA551759	Sequence.	AA551759	92322011
2	AA551759	495 bp	MRNA	EST	11-AUG-1997	11-AUG-1997	2	AA551759	Sequence.	AA551759	92322011
3	AA551759	495 bp	MRNA	EST	11-AUG-1997	11-AUG-1997	3	AA551759	Sequence.	AA551759	92322011
4	AA551759	495 bp	MRNA	EST	11-AUG-1997	11-AUG-1997	4	AA551759	Sequence.	AA551759	92322011
5	AA551759	495 bp	MRNA	EST	11-AUG-1997	11-AUG-1997	5	AA551759	Sequence.	AA551759	92322011
6	AA551759	495 bp	MRNA	EST	11-AUG-1997	11-AUG-1997	6	AA551759	Sequence.	AA551759	92322011
7	AA551759	495 bp	MRNA	EST	11-AUG-1997	11-AUG-1997	7	AA551759	Sequence.	AA551759	92322011
8	AA551759	495 bp	MRNA	EST	11-AUG-1997	11-AUG-1997	8	AA551759	Sequence.	AA551759	92322011
9	AA551759	495 bp	MRNA	EST	11-AUG-1997	11-AUG-1997	9	AA551759	Sequence.	AA551759	92322011
10	AA551759	495 bp	MRNA	EST	11-AUG-1997	11-AUG-1997	10	AA551759	Sequence.	AA551759	92322011
11	AA551759	495 bp	MRNA	EST	11-AUG-1997	11-AUG-1997	11	AA551759	Sequence.	AA551759	92322011
12	AA551759	495 bp	MRNA	EST	11-AUG-1997	11-AUG-1997	12	AA551759	Sequence.	AA551759	92322011
13	AA551759	495 bp	MRNA	EST	11-AUG-1997	11-AUG-1997	13	AA551759	Sequence.	AA551759	92322011
14	AA551759	495 bp	MRNA	EST	11-AUG-1997	11-AUG-1997	14	AA551759	Sequence.	AA551759	92322011
15	AA551759	495 bp	MRNA	EST	11-AUG-1997	11-AUG-1997	15	AA551759	Sequence.	AA551759	92322011
16	AA551759	495 bp	MRNA	EST	11-AUG-1997	11-AUG-1997	16	AA551759	Sequence.	AA551759	92322011
17	AA551759	495 bp	MRNA	EST	11-AUG-1997	11-AUG-1997	17	AA551759	Sequence.	AA551759	92322011
18	AA551759	495 bp	MRNA	EST	11-AUG-1997	11-AUG-1997	18	AA551759	Sequence.	AA551759	92322011
19	AA551759	495 bp	MRNA	EST	11-AUG-1997	11-AUG-1997	19	AA551759	Sequence.	AA551759	92322011
20	AA551759	495 bp	MRNA	EST	11-AUG-1997	11-AUG-1997	20	AA551759	Sequence.	AA551759	92322011
21	AA551759	495 bp	MRNA	EST	11-AUG-1997	11-AUG-1997	21	AA551759	Sequence.	AA551759	92322011
22	AA551759	495 bp	MRNA	EST	11-AUG-1997	11-AUG-1997	22	AA551759	Sequence.	AA551759	92322011
23	AA551759	495 bp	MRNA	EST	11-AUG-1997	11-AUG-1997	23	AA551759	Sequence.	AA551759	92322011
24	AA551759	495 bp	MRNA	EST	11-AUG-1997	11-AUG-1997	24	AA551759	Sequence.	AA551759	92322011
25	AA551759	495 bp	MRNA	EST	11-AUG-1997	11-AUG-1997	25	AA551759	Sequence.	AA551759	92322011
26	AA551759	495 bp	MRNA	EST	11-AUG-1997	11-AUG-1997	26	AA551759	Sequence.	AA551759	92322011
27	AA551759	495 bp	MRNA	EST	11-AUG-1997	11-AUG-1997	27	AA551759	Sequence.	AA551759	92322011
28	AA551759	495 bp	MRNA	EST	11-AUG-1997	11-AUG-1997	28	AA551759	Sequence.	AA551759	92322011
29	AA551759	495 bp	MRNA	EST	11-AUG-1997	11-AUG-1997	29	AA551759	Sequence.	AA551759	92322011
30	AA551759	495 bp	MRNA	EST	11-AUG-1997	11-AUG-1997	30	AA551759	Sequence.	AA551759	92322011
31	AA551759	495 bp	MRNA	EST	11-AUG-1997	11-AUG-1997	31	AA551759	Sequence.	AA551759	92322011
32	AA551759	495 bp	MRNA	EST	11-AUG-1997	11-AUG-1997	32	AA551759	Sequence.	AA551759	92322011
33	AA551759	495 bp	MRNA	EST	11-AUG-1997	11-AUG-1997	33	AA551759	Sequence.	AA551759	92322011
34	AA551759	495 bp	MRNA	EST	11-AUG-1997	11-AUG-1997	34	AA551759	Sequence.	AA551759	92322011
35	AA551759	495 bp	MRNA	EST	11-AUG-1997	11-AUG-1997	35	AA551759	Sequence.	AA551759	92322011
36	AA551759	495 bp	MRNA	EST	11-AUG-1997	11-AUG-1997	36	AA551759	Sequence.	AA551759	92322011
37	AA551759	495 bp	MRNA	EST	11-AUG-1997	11-AUG-1997	37	AA551759	Sequence.	AA551759	92322011
38	AA551759	495 bp	MRNA	EST	11-AUG-1997	11-AUG-1997	38	AA551759	Sequence.	AA551759	92322011
39	AA551759	495 bp	MRNA	EST	11-AUG-1997	11-AUG-1997	39	AA551759	Sequence.	AA551759	92322011
40	AA551759	495 bp	MRNA	EST	11-AUG-1997	11-AUG-1997	40	AA551759	Sequence.	AA551759	92322011
41	AA551759	495 bp	MRNA	EST	11-AUG-1997	11-AUG-1997	41	AA551759	Sequence.	AA551759	92322011
42	AA551759	495 bp	MRNA	EST	11-AUG-1997	11-AUG-1997	42	AA551759	Sequence.	AA551759	92322011
43	AA551759	495 bp	MRNA	EST	11-AUG-1997	11-AUG-1997	43	AA551759	Sequence.	AA551759	92322011
44	AA551759	495 bp	MRNA	EST	11-AUG-1997	11-AUG-1997	44	AA551759	Sequence.	AA551759	92322011
45	AA551759	495 bp	MRNA	EST	11-AUG-1997	11-AUG-1997	45	AA551759	Sequence.	AA551759	92322011

ALIGNMENTS

1 (bases 1 to 495)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1350
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Elias Campo, M.D., Michael R. Emmert-Buck,
M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arraying: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/dbp/imap/image.html

Seq primer: -40m13 fwd. RT from Amersham
High quality sequence stop: 385.
Location/Qualifiers
1. 495
/organism="Homo sapiens"
/notes="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA

was prepared from 12 pooled bulk tumor samples and primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization.

/db_xref="taxon:9606"
/clone="IMAGE:928032"
/clone_1lb="NCI-CGAP_C03"
/sex="pooled"
/tissue="colon"
/tissue_type="DH10B"
/lab_host="DH10B"
<1. ->495

BASE COUNT 89 a 145 c 162 g 99 t
ORIGIN

Query Match 19.6%; Score 470; DB 8; Length 495;
Best Local Similarity 98.8%; Pred. No. 0.00e+00;
Matches 491; Conservative 0; Mismatches 3; Indels 3; Gaps 3;

Db 1 TTCCCTGAAATGGCTTATTTCCAGTGTGAGAAATCTGTGAGTGTGAGGCTCTGT 60
CP 2352 TTCCCTGAAATGGCTTATTTCCAGTGTGAGAAATCTGTGAGTGTGAGGCTCTGT 2293
Db 61 GGTCCCTGTACCCCAAGAGACCTGACCTCCAGATGATTAAGCGCCGGGATCCAGC 120
CP 2292 GGTCCCTGTACCCCAAGAGACCTGACCTCCAGATGATTAAGCGCCGGGATCCAGC 2233
Db 121 CTGTATCTCTCCAGAGCTGGAGCTGTGTGGAGAGCGCGGATGCTGTGTGTAG 180
CP 2232 CTGTATCTCTCCAGAGCTGGAGCTGTGTGGAGAGCGCGGATGCTGTGTGTAG 2173
Db 181 GATGACATCCCAAGGTGGTCTGACAGTGGCCAGATGAGCATGGGCTCACCCTCAG 240
CP 2172 GATGACATCCCAAGGTGGTCTGACAGTGGCCAGATGAGCATGGGCTCACCCTCAG 2113
Db 241 GACAAGGCCACAGGTGGGCGGCGGAGCCAGATGAGCTTACTCTAGAGCAAAATC 300
CP 2112 GACAAGGCCACAGGTGGGCGGCGGAGCCAGATGAGCTTACTCTAGAGCAAAATC 2053
Db 301 CCTCTGTGCTTCTCTTGAAGTCCGCAAGAGGCTCAGTCTTTGAGACCCAGGAGG 360
CP 2052 CCTCTGTGCTTCTCTTGAAGTCCGCAAGAGGCTCAGTCTTTGAGACCCAGGAGG 1993
Db 361 TCAGGGGTGGGCGGCGGCGGAGCCAGATGAGCTTACTCTAGAGCAAAATC 417
CP 1992 TCAGGGGTGGGCGGCGGCGGAGCCAGATGAGCTTACTCTAGAGCAAAATC 1933
Db 418 CACCCAGAGCGGCTACACTGTGAGCTCCGCTCCAGATTTTCAAGGCGCTGTTCG 477
CP 1932 CACCCAGAGCGGCTACACTGTGAGCTCCGCTCCAGATTTTCAAGGCGCTGTTCG 1873
Db 478 TACTCGCGGATGTGTCC 494
CP 1872 TACTCGCGGATGTGTCC 1856

RESULT 2
LOCUS AA633315 441 bp mRNA
DEFINITION ng58f08.s1 NCI-CGAP_C09 Homo sapiens cDNA clone IMAGE:1148103, mRNA
ACCESSION AA633315
NID 92556729
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryote; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 441)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncigap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)

COMMENT

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Ian Kirsch, M.D., Michael R. Emmert-Buck,
M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LINL at:
www-bio.lnl.gov/dbp/image/image.html

Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 433.
Location/Qualifiers

FEATURES

source 1.441
/organism="Homo sapiens"
/note="Organ: colon; Vector: pT73D-Pac (Pharmacia) with a
modified polylinker; 1st strand cDNA was prepared from
RER+ colon tumor, and was then primed with a Not I -
oligo(dT) primer. Double-stranded cDNA was ligated to Eco
RI adaptors (Pharmacia), digested with Not I and cloned
into the Not I and Eco RI sites of the modified pT73
vector. Library is not normalized. Library was
constructed by Bento Soares and M. Fatima Bonaldo
(Soares4)."
/db_xref="taxon:9606"
/clone="IMAGE:1148103"
/clone_1lb="NCI-CGAP_C09"
/tissue_type="Colon tumor RER+"
/lab_host="DH10B"

BASE COUNT 84 a 126 c 144 g 87 t
ORIGIN

Query Match 17.3%; Score 414; DB 10; Length 441;
Best Local Similarity 98.0%; Pred. No. 0.00e+00;
Matches 433; Conservative 0; Mismatches 7; Indels 2; Gaps 2;

Db 1 TTTTCTGGAATGGCTTATTTCCAGTGTGAGAAATCTGTGAGTGTGAGGCTCTGT 60
CP 2355 TTTTCTGGAATGGCTTATTTCCAGTGTGAGAAATCTGTGAGTGTGAGGCTCTGT 2296
Db 61 TGTGTCTCTGTATCCCAAGAGACCTGACAGCTCCAGATGATTAAGCGCCGGATCCA 120
CP 2295 TGTGTCTCTGTATCCCAAGAGACCTGACAGCTCCAGATGATTAAGCGCCGGATCCA 2236
Db 121 GGCCTGTATCTCTCCAGAGCTGGAGCTGTCTAGAGAGCGCGGATCTGTGTGT 180
CP 2235 GGCCTGTATCTCTCCAGAGCTGGAGCTGTCTAGAGAGCGCGGATCTGTGTGT 2176
Db 181 AAGGATGACATCCCAAGAGTGTCTGACAGTGGCCAGATGAGATGGGCTCACCTC 240
CP 2175 AAGGATGACATCCCAAGAGTGTCTGACAGTGGCCAGATGAGATGGGCTCACCTC 2116
Db 241 AAGGACAGGCCACAGTGTGGGGGCGGAGCCAGATGAGCTTACTCTAGAGCAAA 300
CP 2115 AAGGACAGGCCACAGTGTGGGGGCGGAGCCAGATGAGCTTACTCTAGAGCAAA 2056
Db 301 ATCCCTGTGGGCAATCTCTTGAAGTCCGCGCAGAGGCTCAGTCTTTGAGACCCGCG 360
CP 2055 ATCCCTGTGGGCGGCTTCTCTTGAAGTCCGCGCAGAGGCTCAGTCTTTGAGACCCGCG 1996
Db 361 AGGTCCAGGGGTGGGCGGCGGCGGAGCCAGAGGCGGATGAGAGCGGCTGAGGCTCGCG 420
CP 1995 AGGTCCAGGGGTGGGCGGCGGCGGAGCCAGAGGCGGATGAGAGCGGCTGAGGCTCGCG 1937
Db 421 -ACCAACCCAGAGAGCGGCTA 441
CP 1936 CACCAACCCAGAGAGCGGCTA 1915

RESULT 3

LOCUS	AA694490	444 bp	MRNA	EST	08-JAN-1998
DEFINITION	h3h3h08.s1 Soares testis NHT Homo sapiens cDNA clone 1276383 3'				
ACCESSION	AA694490				
NID	92695428				
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryote; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;				
AUTHORS	Primates; Catarrhini; Homnidae; Homo.				
TITLE	1 (bases 1 to 444)				
JOURNAL	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .				
COMMENT	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index Unpublished (1997)				
CONTACT	Contact: Robert Strausberg, Ph.D.				
TEL	Tel.: (301) 496-1550				
EMAIL	Email: Robert.Strausberg@nih.gov				
CDNA LIBRARY PREPARATION	CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima				
BONALDO, Ph.D.	CDNA Library Arrayed by: Greg Lannon, Ph.D.				
DNA SEQUENCING BY	DNA sequencing by: Washington University Genome Sequencing Center				
CLONE DISTRIBUTION	Clone distribution: NCI-CGAP clone distribution information can be				
FOUND THROUGH THE I.M.A.G.E. CONSORTIUM/HLN at:	www.bio.lnl.gov/db/rrp/image/image.html				
Seq primer:	-40m13 fwd, ET from Amersham.				
Location/Qualifiers	1..444				
/organism="Homo sapiens"					
/host="vector: p773b-Pac (Pharmacia) with a modified					
polylinker: Site_1: Not I; Site_2: Eco RI; 1st strand cDNA					
was prepared from mRNA obtained from Clontech					
laboratories, Inc., and primed with a Not I - oligo(dt)					
primer [5',					
TGTTACCAATCTGAAGTGGAGGCGGCCGCCCAATTTTTTTTTTTT 3'].					
Double-stranded cDNA was ligated to Eco RI adaptors					
(Pharmacia), digested with Not I and cloned into the Not I					
and Eco RI sites of the modified p773b vector. Library					
went through one round of normalization to Cots, and was					
constructed by Bento Soares and M. Fatima Bonaldo."					
/db_xref="taxon:9606"					
/clone="1276383"					
/clone_lib="Soares testis NHT"					
/sex="male"					
/lab_host="DH10B"					
BASE COUNT	84 a 125 c 149 g 86 t				
ORIGIN					
Query Match	16.8%; Score 402; DB 11; Length 444;				
Best Local Similarity	96.2%; Pred. No. 0.00e+00;				
Matches 429; Conservative	0; Mismatches 15; Indels 2; Gaps 2;				
1 CCTCTGAAATGGCTTATTTCCTCCACAGTGTGAGGAATCTGTGAGTGTGAGGAGGTCTGTGG	60				
2350 CCTCTGAAATGGCTTATTTCCTCCACAGTGTGAGGAATCTGTGAGTGTGAGGAGGTCTGTGG	2291				
61 TCCCTGTTACCCCAAGAGACCTGTGACCTCCAGATGATTAACGGCCGGGATCCAGAGGCT	120				
2290 TCCCTGTTACCCCAAGAGACCTGTGACCTCCAGATGATTAACGGCCGGGATCCAGAGGCT	2231				
121 TGATCTCTCCAGGCTGGGACTGTGTCTGGGAGAGAGCCGGGCATGCTGTGTGTAAGA	180				
2230 TGATCTCTCCAGGCTGGGACTGTGTCTGGGAGAGAGCCGGGCATGCTGTGTGTAAGA	2171				
181 TGAACATCCCAAGAGTGGTCTGTGACATGTGGCCCAAGATGAGATGGGGCTCACTCAAGA	240				
2170 TGAACATCCCAAGAGTGGTCTGTGACATGTGGCCCAAGATGAGATGGGGCTCACTCAAGA	2111				
241 CAAGGCACACAGTGTGGGGGGCGCATGGACATATGAGGCTTACTCTGGAGCAAAATCCC	300				
2110 CAAGGCACACAGTGTGGGGGGCGCATGGACATATGAGGCTTACTCTGGAGCAAAATCCC	2051				

Db 301 CAGTGGGCTGTCTCTCTGAACTCCGACGAGGCGCTAGCTTTTGAGACCCAGCAGGTC 360
CP 2050 CTGTGGGGGGCTTCTCTCTGAACTCCGACGAGGCGCTAGCTTTTGAGACCCAGCAGGTC 1991
Db 361 AGGGGGTGGCGGCC-ACCTGGGGGAGCATAGGACAGAGCGGCTCAGAGCGCTCGGC-ACCCA 418
CP 1990 AGGGGGTGGCGGCCACCTGGGGGAGCATAGGACAGAGCGGCTCAGAGCGCTCGGCACCCA 1931
Db 419 CCCGAGAGCGGCGCTACACTGCTGA 444
CP 1930 CCCGAGAGCGGCGCTACACTGCTGA 1905

RESULT 4
LOCUS AA708532 400 bp mRNA EST 24-DEC-1997
DEFINITION Z133d12.91 Soares pregnant uterus NDHPU Homo sapiens cDNA clone
506615.3, mRNA sequence.
ACCESSION AA708532
NID 92718450
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 400)
Haller, L., Allen, M., Bowles, L., Dubuque, T., Gelsel, G., Joet, S.,
Kizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Matra, M.,
Matlin, J., Moore, B., Schellenberg, K., Stepice, M., Tan, F.,
Thaisang, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.
MashU-NCI human EST Project
Unpublished (1997)

TITLE
JOURNAL
COMMENT

CONTACT: Wilson R.
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: eastwats@wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Seq primer: -40m13 fwd, 5' from Amersham
High quality sequence stop: 389.
Location/Qualifiers
1. 400
/organism="Homo sapiens"
/note="Organ: uterus; Vector: p7T73-Pac; Site: 1: Not I;
Site: 2: Eco RI; 1st strand cDNA was primed with a Not I -
oligo(dT) primer [5',
ACCTGAGAGATCCGCGCGCGCCCTTTTCTTTTCTTTTCTTTT 3']
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
went through one round of normalization. Library
constructed by M. Fatima Boudado.
/db_xref="GDB:3812634"
/db_xref="taxon:9606"
/clone="506615"
/clone_id="Soares pregnant uterus NDHPU"
/sex="female"
/dev_stage="adult"
/lab_host="DH10s"
/lab="107 c 136 g 81 t

BASE COUNT 76 a 107 c 136 g 81 t
ORIGIN

Query Match 16.5% Score 396; DB 11; Length 400;
Best Local Similarity 99.5%; Pred. No. 0.00e+00;
Matches 398; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 1 TTTTCCCTGTAATGCGTTTATTTCCCAAGTGTGAGGAATCTGTGAGTGTGAGGCTCT 60
TTTTCCCTGTAATGCGTTTATTTCCCAAGTGTGAGGAATCTGTGAGTGTGAGGCTCT 2295

Db 61 GTGTCCTCCTGTTACCCCAAGACCTGAGCTCCAGATGATAGGCGCGGATCCAG 120
 Cp 2294 GTGGTCCTCCTGTTACCCCAAGACCTGAGCTCCAGATGATAGGCGCGGATCCAG 2235
 Db 121 GCTTCGATCCTCCAGAGCTGGAGTGTCTGGAGAGAGCGGGCATCTGTGTTGTA 180
 Cp 2234 GCTTCGATCCTCCAGAGCTGGAGTGTCTGGAGAGAGCGGGCATCTGTGTTGTA 2175
 Db 181 AGGATGACACCTCCAAAGGTGTCTGTACATGAGCCCGCATGATGAGGCTCACTCA 240
 Cp 2174 AGGATGACACCTCCAAAGGTGTCTGTACATGAGCCCGCATGATGAGGCTCACTCA 2115
 Db 241 AGGACAGAGCCCAAGAGTGGCGGGCCAGAGCCAGATGAGCTTACTCTAGAGCAAAA 300
 Cp 2114 AGGACAGAGCCCAAGAGTGGCGGGCCAGAGCCAGATGAGCTTACTCTAGAGCAAAA 2055
 Db 301 TCCCTGTGGGGGCTTCTCTTGAAGTCCGCGCAGAGGCTCAGTCTTTGAGCCAGCA 360
 Cp 2054 TCCCTGTGGGGGCTTCTCTTGAAGTCCGCGCAGAGGCTCAGTCTTTGAGCCAGCA 1955
 Db 361 GGTACGGGGGTGGGGGCGCCACCTGGGGGCGAGCAAGCCAGA 400
 Cp 1954 GGTACGGGGGTGGGGGCGCCACCTGGGGGCGAGCAAGCCAGA 1955

RESULT 5
 LOCUS AA878567 381 bp mRNA EST 25-MAR-1998
 DEFINITION 0117D06.s1 NCI_CGAP_K1d5 Homo sapiens cDNA clone IMAGE:1492403 3',
 mRNA sequence.
 ACCESSION AA878567
 NID 92987532
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
 Primates; Catarrhini; Hominoidea; Homo.
 1 (bases 1 to 381)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert.Strausberg@nih.gov
 Tissue Procurement: Christopher Koskaju, M.D., Ph.D., Michael R.
 Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D.
 DNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 www.bio.llnl.gov/dbp/image/image.html

Seq primer: -40m13 fwd. ET from Amerisham
 High quality sequence stop: 381.
 Location/Qualifiers
 1. 381

/organism="Homo sapiens"
 /note="Organ: Kidney; Vector: pT73D-Pac (Pharmacia) with
 a modified polylinker. Site: 1. Not I; Site 2. Eco RI; 1st
 strand cDNA was primed with a Not I - oligo(dT) primer [5'-
 AACTGAGAAATTCGCGCGCGCAATATTTTATTTTATTTT-3']
 double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT73 vector. Library
 went through one round of normalization. Library
 constructed by Bento Soares and M. Fatima Bonaldo.
 /db_xref="taxon:9606"
 /clone_image:1492403
 /clone_11b="NCI_CGAP_K1d5"
 /tissue_type="7 pooled tumors (clear cell type)"
 /lab_host="DH10B"

BASE COUNT 70 a 101 c 126 g 84 t
 ORIGIN
 Query Match 15.8%; Score 377; DB 13; Length 381;
 Best Local Similarity 99.5%; Pred. No. 0.00e+00;
 Matches 379; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Db 1 TTTTTCCTCTGAAATGCTTATTTCCAGATGAGGATCTGTGTAGG 60
 Cp 2358 TTTTTCCTCTGAAATGCTTATTTCCAGATGAGGATCTGTGTAGG 2299
 Db 61 GTCTGTGTCCTGTTACCCCAAGAGCCTGACCTCCAGATGATTAAGCGCGGAT 120
 Cp 2298 GTCTGTGTCCTGTTACCCCAAGAGCCTGACCTCCAGATGATTAAGCGCGGAT 2239
 Db 121 CCAGGCTTGAATCTCCAGAGCTGGAGACTGTCTGGAGAGAGCCGGCATCTGTGTT 180
 Cp 2238 CCAGGCTTGAATCTCCAGAGCTGGAGACTGTCTGGAGAGAGCCGGCATCTGTGTT 2179
 Db 181 TGTAGGATGACACTCCCAAGGTGTCTGTACAGTGGCCAGATGAGATGGGCTCAC 240
 Cp 2178 TGTAGGATGACACTCCCAAGGTGTCTGTACAGTGGCCAGATGAGATGGGCTCAC 2119
 Db 241 CTCAGAGACAGGCCACAGAGTGGCGGGCCAGAGCCAGATGAGCTTACTCTAGAGC 300
 Cp 2118 CTCAGAGACAGGCCACAGAGTGGCGGGCCAGAGCCAGATGAGCTTACTCTAGAGC 2059
 Db 301 AAAATCCCTGTGGGGGCTTCTCTTGAAGTCCGCGCAGAGGCTCAGTCTTTGAGCCCA 360
 Cp 2058 AAAATCCCTGTGGGGGCTTCTCTTGAAGTCCGCGCAGAGGCTCAGTCTTTGAGCCCA 1999
 Db 361 GGCAGGTCAAGGGGTGGCGGC 381
 Cp 1998 GGCAGGTCAAGGGGTGGCGGC 1978

RESULT 6
 LOCUS N31660 413 bp mRNA EST 10-JAN-1996
 DEFINITION YX69D04.r1 Homo sapiens cDNA clone 266959 5'.
 ACCESSION N31660
 NID 91152059
 KEYWORDS EST.
 SOURCE human clone-266959 primer-T7 library-Soares melanocyte 2BHM
 vector-pT73D (Pharmacia) with a modified polylinker host-DH10B
 (ampicillin resistant) Raltei-Not I Raltei-Eco RI Male. 1st strand
 cDNA was primed with a Not I - oligo(dT) primer
 [5'-TGTACCAATCTGAAAGTGGAGCGCGCGCATTTTATTTTATTTT-3']
 double-stranded cDNA was size selected, ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I and Eco
 RI sites of a modified pT73 vector (Pharmacia). Library
 constructed by Bento Soares and M. Fatima Bonaldo. RNA from normal
 foreskin melanocytes (FS374) was kindly provided by Dr. Anthony P.
 Albino.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
 Primates; Catarrhini; Hominoidea; Homo.
 1 (bases 1 to 413)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished (1995)
 COMMENT
 Contact: Wilson RK
 Washu-Merck EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810

Email: estewartson.wustl.edu
High quality sequence crops: 294
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Location/Qualifiers
1 413

Source	1.	.413	/Organism="Homo sapiens"	2
mRNA	<1.	.>413	/clone="266959"	2
BASE COUNT	82 a	124 c	132 g	73 t
ORIGIN	2 others			

Query Match	15.8%	Score 379;	DB 19;	Length 413;
Best Local Similarity	97.3%;	Pred. No. 0.00e+00;		
Matches	403;	Conservative	0;	Mismatches 8;
			Indels	3;
			Gaps	3

Db 1 TAGGACAAAGCGGGAGAGCCGACTCCGAGCTGTGAAGCGGCACGTCGCCAGAAAGGTGGACTT 60
 |||||
Oy 1780 TAGGACAAGCGGGAGAGCGACTCCGAGGCTCTGAAGCGGCACGTCGCCAGAAAGGTGGACTT 180

Db 61 GGCACTGAAACAGCTGGGACACATCCGCGAGTACGAAACAGGCGCTGAAAGTCTGGAGCG 12
 |||||
 0y 1840 GGCACTGAAACAGCTGGGACACATCCGCGAGTACGAAACAGGCGCTGAAAGTCTGGAGCG 18

D_b 121 GGAGGTCCA-CAGTGTAGCCGCGCTCTGGGGTGGGTCGACAGGCC-TGAGCCGCTCTGC 17
 |||||
O_y 1900 GGAGGTCCACAGTGTAGCCGCGCTCTGGGGTGGGTCGACAGGCCCTGAGCCGCTCTGC 19

D_b 179 CTTCCTGCCCCCAGGTGGGCCCGCCACCCCTGACCTGCCCTGGGTCCAAAGACTGAGCCT 23
 |||||
 1960 CTTCCTGCCCCCAGGTGGGCCCGCCACCCCTGACCTGCCCTGGGTCCAAAGACTGAGCCT 20

Db 239 GCTGGCGACTTCAAGAGGAAGCCCCACAGGAAATTTTGTCTAGAGTAAGGCTCATC 29
|||||
|||
Oy 2020 GCTGGCGACTTCAAGAGGAAGCCCCACAGGGAATTTTGTCTCTAGAGTAAGGCTCATC 20
|||||
|||

Db 299 TGGCCCTCGGCACGGCACCCTGGTGGCTTGTCTCTGAGGTGAGCCCCATGTCCTCG 35
|||||
Qy 2080 TGGCCCTCGGCCCCCGCACCCTGGTGGCTTGTCTCTGAGGTGAGCCCCATGTCCTCG 21

Db 359 GCCACTGTGAGGACAACTTTGGGAGTGTATCTTACAAANCAAGCATGCC 412
|||||
Oy 2140 GCCACTGTGAGGACCACTTTGGGA-GTGTCAATCTTACAAACCAAGCATGCC 2192
|||||

RESULT	7	H18835	410 bp	mRNA	EST	29-JUN-1995

ACCESSION
H18835
NID
9885075
KEYWORDS
EST.

strand cDNA was primed with a Not I - oligo(dT) primer [5,

Not I and directionally cloned into the Not I and Hind III sites of the LacZ reporter vector. Library went through one round of

ORGANISM Homo sapiens
Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
Bonnado.
Formulation: library constructed by Bento Soares and M.Falim.

REFERENCE
AUTHORS
1 (bases 1 to 410)
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawking, M.,
Eutheria; Archonta; Primates; Carnivora; Hominoidea; Homo.

Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,
Trevaskis, E., Waterston, R., Williamson, A., Wohlmann, P. and
Wilson, R.

TITLE	JOURNAL	COMMENT
The WashU-Merck EST Project	Unpublished (1995)	

GDB: G00-474--072
Contact: Wilson RK
Washington-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estevenson.wustl.edu

Email: est@station.wustl.edu
 High quality sequence stops: 367
 Source: IMAGE Consortium, LNL
 This clone is available royalty-free through LNL ; contact the
 IMAGE Consortium (info@image.jnl.gov) for further information.
 Location/Qualifiers :

	BASE COUNT	ORIGIN
/organism="Homo sapiens"	59 a	130 c
/clone="51262"	127 g	90 t
		4 others

Query Match	14.6%	Score 350;	DB 16;	Length 410;
Best Local Similarity	94.6%	Pred. No. 7,30e-292;		
Matches	384;	Conservative	0;	Mismatches 19;
			Indels 3;	Gaps 3

Db **1** TGCAGNGAACCTTGAACGGGAGGTGTGCGCAGCCCTCGCCAGCGGGGCCCCCGGCGTGNGC 60
 ||||| | ||||| ||||| | | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Gy 856 TGCGCCAGGGCCTGAACGGGAGGGGCG -GACACTTCGCCAAGCGGGGGCCCCCGGCGCTTGGC 914

Db 61 CATTGCCCTACTGAGCCAGCGGGCTGC-CTNCTACCTCGCCGACAGCTTGGAAACCAGTGCGGA 119
 ||| |||||||
Gy 915 CAT-GCCTCACTGAGCCAGCGGGCTTCACCTCGCCGACAGCTTGGAAACCAGTGCGGA 973

Ddb 120 CCGTAGTGGGCTCAACGTGCCTTCCTCCTGGGCGATGGGCTCCGGGCTGAACCCCAGGTTTGTA 179

QY 974 CCGTAGTGGGCTCAACGTGCCTTCCTCCTGGGCGATGGGCTCCGGGCTGAACCCCAGGTTTGTA 1033

Db 180 CCACCTGGGGCCGACACTGTCCCTCTGCATCGACTTCATGGTTTTCAACGGTGCGGCTGCTTCA 239
|||||
QY 1034 CCACCTGGGGCCGACACTGTCCCTCTGCATCGACTTCATGGTTTTCAACGGTGCGGCTGCTTCA 1099

Db 240 CATCTTCACGGTCAACAACAGCTGGGCCCAAGATCGTCATCGAGCAAGATGATGAA 299
|||||
QY 1094 CATCTTCACGGTCAACAACAGCTGGGCCCAAGATCGTCATCGTCGAGCAAGATGATGAA 1155

Db	300	GGACGTGTTCTTCTTCCCTTCTTCCGCGGCGGTGAGCTATGGGTTGGGCCAC	359
QY	1154	GGACGTGTTCTTCTTCCCTTCTTCCGCGGCGGTGAGCTATGGCGTGGCCAC	1211

Db			
360	GGAGGGGCTTCTGAGGCCACGGGACAGTGCATTCCCAAGTATCTTG	405	
QY	1214 GGAGGGGCTTCTGAGGCCACGGGACAGTGCATTCCCAAGTATCTTG	1259	

RESULT	8			
LOCUS	AI092908	311 bp	mRNA	EST
				18-AUG-1998

3' / mRNA sequence.
ACCESSION A1092908
NTD 93431884

SOURCE ORGANISM	
human.	
Homo sapiens	
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;	

REFERENCE	AUTHORS	TITLE
1 (bases 1 to 311)	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .	National Cancer Institute, Cancer Genome Anatomy Project (CGAP)

JOURNAL
UNIVERSITY OF TORONTO LIBRARY
UNPUBLISHED (1997)

COMMENT

Contact: Robert Strausberg, Ph.D:

Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov

This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.lnl.gov) for further information.
Seq primer: -40m3 fwd. RT from Amersham
High quality sequence stop: 272.

FEATURES

Location/Qualifiers
1. 311

/organism="Homo sapiens"
/note="Organ: mixed (see below); Vector: p773D-Pac (Pharmacia) with a modified polylinker; Site.1: Not I; Site.2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (melanocyte 2NBH, pregnant uterus NBHPU, and fetal heart NBH19H) were mixed, and as circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 260232-265223, 340488-345479, and 484488-489479."
/db_xref="taxon:9606"
/clone="IMAGE:1688836"
/clone_lib="Soares_NHMPU_S1"
/tissue_type="Pooled human melanocyte, fetal heart, and pregnant uterus"
/lab_host="DH10B"

BASE COUNT 65 a 81 c 93 g 72 t
ORIGIN

Query Match 12.4%; Score 301; DB 17; Length 311;
Best Local Similarity 98.4%; Pred. No. 7.78e-245;

Matches 306; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Db 1 TTTTCTCTGCTGAAATGGCTTTATTTCCCGAGTGTGAGGAATCTGTAGTGTGAGG 60
Cp 2359 TTTTCTCTGCTGAAATGGCTTTATTTCCCGAGTGTGAGGAATCTGTAGTGTGAGG 2360
Db 61 GGTCTGTGCTCTGTTACCCCAAGAGACCTGACGCTCCCAATGGATACGGCCCGGA 120
Cp 2239 GGTCTGTGCTCTGTTACCCCAAGAGACCTGACGCTCCCAATGGATACGGCCCGGA 2240
Db 121 TCCAGGCTTGTATCTCTCCCAAGCTGGAGTGTCTGTGGAGAGCCGGCATGCTGTGT 180
Cp 2239 TCCAGGCTTGTATCTCTCCCAAGCTGGAGTGTCTGTGGAGAGCCGGCATGCTGTGT 2180
Db 181 TTGTAAGATGACACTCCCAAGGTGTCTGTACAGTGTGCGCCCAATGAGATGGGGCTCA 240
Cp 2179 TTGTAAGATGACACTCCCAAGGTGTCTGTACAGTGTGCGCCCAATGAGATGGGGCTCA 2120
Db 241 CCTCAAGACAGGCCACAGTGTGGGGGGCCGAGGCCCAATGAGATGAGCTTACTTTAGAG 300
Cp 2119 CCTCAAGACAGGCCACAGTGTGGGGGGCCGAGGCCCAATGAGATGAGCTTACTTTAGAG 2060
Db 301 CAAAATCCCT 311
Cp 2059 CAAAATCCCT 2049

RESULT 9
LOCUS AA932133 309 bp mRNA
DEFINITION c90h09.s1 NCI-CCAP_K1d3 Homo sapiens cDNA clone IMAGE:1554497 3'
ACCESSION AA932133
NID 93086446
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 309)
AUTHORS NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL COMMENT

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550

Email: Robert.Strausberg@nih.gov

Tissue Procurement: Christopher Koskajuk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: M. Bento Soares, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.

DNA Sequencing by: Washington University Genome Sequencing Center
clone distribution: NCI-CCAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
www.bio.lnl.gov/db/rt/image/image.html

Insert Length: 394 Std Error: 0.00
Seq primer: -40m3 fwd. RT from Amersham
High quality sequence stop: 309.

FEATURES

Location/Qualifiers
1. 309

/organism="Homo sapiens"
/note="Organ: kidney; Vector: p773D-Pac (Pharmacia) with a modified polylinker; Site.1: Not I; Site.2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer, double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified p773 vector. mRNA of normalization. Library went through one round of normalization. Library constructed by Bento Soares and M. Fatima Bonaldo."
/db_xref="taxon:9606"
/clone="IMAGE:1554497"
/clone_lib="NCI-CCAP_K1d3"
/lab_host="DH10B"

BASE COUNT 61 a 80 c 102 g 66 t
ORIGIN

Query Match 12.4%; Score 297; DB 15; Length 309;
Best Local Similarity 98.1%; Pred. No. 5.22e-241;

Matches 303; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Db 1 TTTCTCTGAAATGGCTTTATTTCCCGAGTGTGAGGAATCTGTAGTGTGAGGCTCG 60
Cp 2353 TTTCTCTGAAATGGCTTTATTTCCCGAGTGTGAGGAATCTGTAGTGTGAGGCTCG 2294
Db 61 TGTCTCTGTTACCCCAAGAGACCTGACGCTCCCAATGGATACGGCCCGGA 120
Cp 2239 TGTCTCTGTTACCCCAAGAGACCTGACGCTCCCAATGGATACGGCCCGGA 2234
Db 121 CCTGATCTCCAGAGCTGGAGCTGTCTGGAGAGACCCGGGATCTGTGTTGTA 180
Cp 2233 CCTGATCTCCAGAGCTGGAGCTGTCTGGAGAGACCCGGGATCTGTGTTGTA 2174
Db 181 GGATGACACTCCCAAGGTGTCTGTACAGTGTGCGCCCAATGAGATGGGGCTCA 240
Cp 2173 GGATGACACTCCCAAGGTGTCTGTACAGTGTGCGCCCAATGAGATGGGGCTCA 2114
Db 241 GGACAGGCCACAGAGTGTGGGGGGCCGAGGCCCAATGAGATGAGCTTACTTTAGAG 300
Cp 2113 GGACAGGCCACAGAGTGTGGGGGGCCGAGGCCCAATGAGATGAGCTTACTTTAGAG 2054
Db 301 CCCCTGTGG 309
Cp 2053 CCCCTGTGG 2045

RESULT 10
LOCUS H18836 271 bp mRNA
DEFINITION ym45d10.s1 Homo sapiens cDNA clone 51262 3'
ACCESSION H18836
NID 9885076
KEYWORDS EST.
SOURCE human clone-51262 library-Soares infant brain INIB vector-Lafield BA host-DH10B (ampicillin resistant) primer-Promega -31m3 Raitel-Not

	REFERENCE	I RefSeq-Hind III Whole brain from a 73 days post natal female. 1st strand cDNA was primed with a Not I - oligo(3') primer [5' AACGCGAAGATTGGCCGCCAGCAATTTTCTTTTTTTT 3']; double-stranded cDNA was ligated to Hind III adaptors (Pharmacia), digested with Not I and directionally cloned into the Not I and Hind III sites of the Lambda BA vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo.
	AUTHORS	Homo sapiens Eukaryotes; Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Chonata; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates; Catarrhini; Hominiid; Homo. 1 (bases 1 to 271) Hillier,L., Clark,N., Dubugue,T., Elliston,K., Hawkins,M., Holman,J., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,M., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevisan,E., Waterston,R., Williamson,A., Wollmann,P. and Wilson,R.
	TITLE	The WashU-Merck EST Project
	JOURNAL	Unpublished (1995)
	COMMENT	GDB: G00-424-072 Contact: Wilson RK WashU-Merck EST Project Washington University School of Medicine 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel.: 314 286 1800 Fax: 314 286 1810 Email: est@wustl.wustl.edu High quality sequence stcps: 162 Source: IMAGE Consortium, LML This clone is available royalty-free through LML ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Location/Qualifiers 1..271 /organism="Homo sapiens" /clone="51262"
	FEATURES	BASE COUNT 50 a 64 c 86 g 67 t 4 others ORIGIN
	source	
	Query Match	10.8%; Score 258; DB 16; Length 271; Best Local Similarity 97.8% Pred. No. 7.99e-204;
	Matches	265; Conservative 0; Mismatches 5; Indels 1; Gaps 1;
Db	1	TTTTTTTTTTCCTGTAATAAGGCTTATTCCCCCAGTGAGAAATCTGTGAGTGTGA 60
Cp	2361	TTTTTTTTTTCCTGTAATAAGGCTTATTCCCCCAGTGAGAAATCTGTGAGTGTGA 2302
Db	61	GGGGTCTGTGCTCCCTGTACCACCAAGAGACCTGCAGCTCCAGATGGATTAAGGCCCCGG 120
Cp	2301	GGGGTCTGTGCTCCCTGTACCACCAAGAGACCTGCAGCTCCAGATGGATTAAGGCCCCGG 2242
Db	121	GATCCAGGCTTGTGATCCTCCAGAGNTGGANTGTTCTTGAGAGAGACCGGGCATGCTGTG 180
Cp	2241	GATCCAGGCTTGTGATCCTCCAGAGGCTGGACTGTTCTTGAGAGAGACCGGGCATGCTGTG 2182
Db	181	GTTGTGAAGATGACACTCCCAAAGGTGGTCTGTACAGTAGTGCCCAATGGAGATGGGGANT 240
Cp	2181	GTTGTGAAGATGACACTCCCAAAGGTGGTCTGTACAGTAGTGCCCAATGGAGATGGGGGCT 2122
Db	241	CACCTNAGAAGACAAGGCCACAGAGTGGG 271
Cp	2121	CACCTCAAGACAGAG-GCACACAGGTGGGG 2092
RESULT	11	
LOCUS	AA026974	294 bp mRNA EST 09-MAY-1997
DEFINITION	XK01610.s1 Soares pregnant uterus NBHPU Homo sapiens cDNA clone	
ACCESSION	AA026974	
NID	91493165	
KEYWORDS	EST.	

SOURCE ORGANISM human.
Eukaryote; mtochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Primates; Carnivora; Homnidae;
Homo.

REFERENCE 1 (bases 1 to 294)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,R., Hawkins,M.,
Holman,M., Hultman,M., Kueba,T., Le,M., Lennon,G., Merit,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Tan,F., Trevaskis,E.,
Waterson,R., Williamson,A., Wohlmann,P. and Wilson,R.
WashU-Merck EST Project
Unpublished (1995)

TITLE JOURNAL COMMENT

Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 799 Std Error: 0.00
Seq primer: -40M13 fwd. from Amersham
High quality sequence stop: 212.
Location/Qualifiers
1..294
/organism="Homo sapiens"
/note="Organ: uterus; Vector: pT73-Pac; Site:1: Not I ;
Site_2: Eco RI; 1st strand cDNA was primed with a Not I -
oligo(dT) primer [5'
AAGCGAGAATGCGGCCGCCCTTTTTTTTTTTTTT 3'] ,
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
went through one round of normalization. Library
constructed by M. Fatima Bonaldo."
/db_xref="taxon:9606"
/clone="469290"
/clone_id="Scars pregnant uterus NBPV"
/sex="female"
/dev_stage="adult"
/lab_host="DH10S"
complement(<1..>294).
/db_xref="GDB:j375523".

BASE COUNT 53 a 70 c 88 g 78 t 5 others

ORIGIN

Query Match 10.7% Score 257; DB 6; Length 294;
Best Local Similarity 98.0%; Pred. No. 7,12e-203;
Matches 287; Conservative 0; Mismatches 1; Indels 5; Gaps 5;

D 1 TTTTTCCTGTGTAATGGCTTATTCCCGATGAGAAAC 60
CP 2373 TTTTTCCTGTGTAATGGCTTATTCCCGATGAGAAAC 60
D 61 TTGAGAGTGAGGGGCTGTGCTCCTGTTACCCCAAGACCCTGCAGCTCAGATGG 120
CP 2313 TTGAGAGTGAGGGGCTGTGCTCCTGTTACCCCAAGACCCTGCAGCTCAGATGG 120
D 121 ATACGCCCCGGAATCCAGGCTTGATCCTCCAGAGCTGGAGTGTGAGAGAGC 180
CP 2253 ATACGCCCCGGAATCCAGGCTTGATCCTCCAGAGCTGGAGACT-GGTTTGGAGAGAGC 195
D 181 CGGCGCATGCTGTGTGTTGAAGATGACACTCCCAAAGGTNGTCTCGACANGTAGGCCA 240
CP 2194 CGGCGCATGCTGTGTGTTGAAGATGACACTCCCAAAGG-GGTCTCGANA-GTGCCCA 2137
D 241 NATGGCATGGGGGCTCACCCAAAGACAANGGCCACACAGTGGGGGGCGGA 293
CP 2136 NATGGCATGGGGGCTCACCCAAAGACAANGGCCACACAGTGGGGGGCGGA 2086

LOCUS	12	AA454774	299 bp	MRNA	EST	06-JUN-1997
DEFINITION		277807.r1 Soares ovary tumor NBOT Homo sapiens CDNA clone 809749				
ACCESSION		5', mRNA sequence.				
NID		AA454774				
KEYWORDS		92177550				
SOURCE		EST.				
ORGANISM		human.				
		Homo sapiens				
		Eukaryote; mitochondrial eukaryotes; Metazoa; Chordata;				
		Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae;				
REFERENCE		1 (Pages 1 to 299)				
AUTHORS		Hallier,L., Allen,M., Bowles,L., Dubuque,T., Gettel,G., Jost,S., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marie,M., Martin,J., Moore,B., Schellenberg,R., Stoops,M., Tan,F., Theising,B., White,Y., Wylie,T., Waterson,R. and Wilson,R.				
TITLE		WashU-Merck EST Project 1997				
JOURNAL		Unpublished (1997)				
COMMENT		Contact: Wilson RK WashU-Merck EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@wustl.wustl.edu This clone is available royalty-free through LML; contact the IMAGE Consortium (info@image.llnl.gov) for further information. The vector to vector length is 504 Seq primer: -28m13 rev2 ET from Amersham High quality sequence stop: 288. Location/Qualifiers 1..299 /organism="Homo sapiens" /note="Organ: ovary; Vector: pT73D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer (5 TGTTCACATCTGAGTGGAGCGCGCGGCTTTTCTTTTCTTTT 3'), double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT73 vector (Pharmacia). Library constructed by Bento Soares and M.Felicia Bonaldo. /db_xref="taxon:9606" /db_xref="clone:809749" /clone.lib="Soares ovary tumor NBOT" /sex="Female" /tissue_type="Ovarian tumor" /lab_host="DH10B (ampicillin resistant)" <1..>299 /db_xref="GDB:6039468"				
BASE COUNT	51 a	95 c	99 g	54 t		
ORIGIN						
Query Match		10.7%	Score 255;	DB 7;	Length 299;	
Best Local Similarity		95.4%	Pred. No. 5,656-201;			
Matches	289;	Conservative 0;	Mismatches 10;	Indels 4;	Gaps 4;	
Db	1	CTGAACAGGTGGACACATCGCGAGTACGAAAGCGCGCGAAGTCTGGAGCGGGAG	60			
Oy	1844	CTGAACAGGTGGACACATCGCGAGTACGAAAGCGCGCGAAGTCTGGAGCGGGAG	1903			
Db	61	GTCAGC-AGTGTAGCGCGCTCTGGGGTGGTGGCCGAGCGCTGAGCGCATCGGCTT	119			
Oy	1904	GTCAGCAGTGTAGCGCGCTCTGGGGTGGTGGCCGAGCGCTGAGCGCATCGGCTT	1963			
Db	120	CTGCCCCCAGGTGGGCGCGCACCC-TCACCTGCTTGGGTCCAAAGATGAGCCCTGCTG	178			
Oy	1964	CTGCCCCCAGGTGGGCGCGCACCCCTACCTGCTTGGGTCCAAAGATGAGCCCTGCTG	2023			
	179	GGGACTCTCAAGAGAGAGAGCC-ACAGG-ATTTGTCTCTCAAGTAAAGCTATCTGGG	236			

QY	2024	GGGAGATTCCAGAGAAAGCCCGCCACAGGGGATTTTGTCTCTAGATGAGCTCATCTGGG	2083
Db	237	CGTCGGCCCCCGACAGCGTGTGGCTTGTCCCTGAGCTAGAGCCCATGTCTATCTGGGCCA	296
QY	2084	CTCTGGCCCCCGACACCTGTGTGGCTTGTCTTGTGAGGTGAGCCCATGTCTCATCTGGGCCA	2143
Db	297	CTG 299	
QY	2144	CTG 2146	
RESULT	13		
LOCUS	AA592910	259 bp	mRNA
DEFINITION	nm01407.61	NCI_CGAP_C09	Homo sapiens cDNA clone IMAGE:1076485, mRNA
ACCESSION	AA592910		
NTD	92408672		
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;		
AUTHORS	Primates; Catarrhini; Homnidae; Homo.		
TITLE	1 (bases 1 to 259)		
JOURNAL	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.		
COMMENT	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),		
	Tumor Gene Index		
	Unpublished (1997)		
	Contact: Robert Strausberg, Ph.D.		
	Tel: (301) 496-1550		
	Email: Robert.Strausberg@nih.gov		
	Tissue Procurement: Ilan Kirsch, M.D., Michael R. Emmert-Buck,		
	M.D., Ph.D.		
	cDNA Library Preparation: M. Bento Soares, Ph.D.		
	cDNA Library Arrayed by: Greg Lennon, Ph.D.		
	DNA Sequencing by: Washington University Genome Sequencing Center		
	Clone distribution: NCI-CGAP clone distribution Information can be		
	found through the I.M.A.G.E. Consortium/LLNL at:		
	www-bio.llnl.gov/dbtp/image/image.html		
	Insert Length: 2323	Std Error: 0.00	
	Seq primer: -40ml3 fwd. Ex from Amersham		
	High quality sequence stop: 240.		
FEATURES	Location/Qualifiers		
SOURCE	1..259		
	/organism="Homo sapiens"		
	/note="Organ: colon; Vector: pT73D-Pac (Pharmacia) with a		
	modified polylinker; 1st strand cDNA was prepared from		
	RER: colon tumor, and was then primed with a Not I -		
	oligo(dT) primer. Double-stranded cDNA was ligated to Eco		
	RI adaptors (Pharmacia), digested with Not I and cloned		
	into the Not I and Eco RI sites of the modified pT73		
	vector. Library is not normalized. Library was		
	constructed by Bento Soares and M. Fatima Bonaldo		
	(Soares4)."		
	/db_xref="taxon:9606"		
	/clone="IMAGE:1076485"		
	/clone_1b="NCI_CGAP_C09"		
	/issue_type="colon tumor RER"		
	/lab_host="DH108"		
BASE COUNT	51 a	66 c	86 g 56 t
ORIGIN			
Query Match	10.4%	Score 249,	DB 10; Length 259;
Best Local Similarity	98.8%	Pred. No. 2,80e-195;	
Matches 257;	Conservative	0; Mismatches 2;	Indels 1; Gaps 1;
Db	1	CGTCGAAATGCGTTTATTCGCCAGGTGATGAGTATGAGTGTGAGCGGTCTGTGG	60
Cp	2350	CGTCGAAATGCGTTTATTCGCCAGGTGATGAGTGTGAGTGTGAGCGGTCTGTGG	2291
Db	61	TCCCTGTATCCCAAGACCTCGAGCTCTCAGATGATTAACGGCTGGGATCCAG-CAT	119

 MPELH
 (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

Mparch_n n.a. - n.a. database search, using Smith-Waterman algorithm
 Run on: Sun Feb 14 19:34:43 1999; Master time 3738.48 Seconds
 Tabular output not generated. 1512.673 Million cell updates/sec

Title: >US-08-946-869-9
 Description: (1-2393) from US08946869.seq
 Perfect Score: 2393
 N.A. Sequence: 1 AAGGCCACCTCGCGCTCTC.....AAAAAAGCGCGCGCGGT 2393
 Comp: TTCGCGTGGAGCGCGAGAG.....TTTTTCCGCGCGCGCGCA

Scoring table: TABLE default
 Gap 6

Mmatch STD : Dbase 0; Query 0

Searched: 602357 seqs, 1181590623 bases x 2

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: emb157
 1:em_ba 2:em_fun 3:em_htg 4:em_hum1 5:em_hum2 6:em_in
 7:em_com 8:em_ov 9:em_ov 10:em_pat 11:em_ph 12:em_pi
 13:em_to 14:em_un 15:em_vl
 Database: genbank110
 16:gb_ba1 17:gb_ba2 18:gb_htg 19:gb_in 20:gb_com 21:gb_ov
 22:gb_pat 23:gb_ph 24:gb_pi1 25:gb_pi2 26:gb_pi1
 27:gb_pi2 28:gb_pi3 29:gb_pi3 30:gb_pi3 31:gb_pi3 32:gb_pi3
 33:gb_un 34:gb_vl

Statistics: Mean 12.262; Variance 12.056; scale 1.017

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
C 1	258	10.8	271	31	G22461	human STR WI-30695.	6.03e-89
C 2	118	4.9	6220	28	AB001535	Homo sapiens mRNA, com	5.59e-31
C 3	86	3.6	78064	18	HSAC000376	*** SEQUENCING IN PROG	2.11e-18
C 4	86	3.6	155074	28	AC003693	Human Chromosome 11p15	2.11e-18
C 5	78	3.3	7218	22	I66494	Sequence 14 from paten	2.39e-15
C 6	54	2.3	7218	22	I66494	Sequence 14 from paten	1.51e-06
C 7	53	2.2	78064	18	HSAC000376	*** SEQUENCING IN PROG	3.39e-06
C 8	45	1.9	148869	18	AC006037	*** SEQUENCING IN PROG	1.86e-03
C 9	45	1.9	193381	18	AC005538	*** SEQUENCING IN PROG	1.86e-03
C 10	40	1.7	1016	25	AF005158	Arabidopsis thaliana M	8.08e-02
C 11	40	1.7	2286	29	AF024620	Mus musculus gamma-aml	8.08e-02
C 12	40	1.7	2335	29	RN030290	Rattus norvegicus gala	8.08e-02
C 13	40	1.7	2865	27	HSL1867	Homo sapiens (subclone	8.08e-02

14	40	1.7	3451	22	E13998	Cyanidium caldarium nr	8.08e-02
15	40	1.7	3519	27	HSL1869	Homo sapiens (subclone	8.08e-02
16	40	1.7	6922	29	NC017793	Mus musculus mRNA for	8.08e-02
17	40	1.7	74371	27	AC005369	Homo sapiens Chromosom	8.08e-02
18	41	1.7	75270	28	AF023268	Homo sapiens ckr2 kina	3.85e-02
19	41	1.7	151239	18	HS377F16	Human DNA sequence ***	3.85e-02
20	38	1.6	965	22	AR024329	Sequence 22 from paten	3.50e-01
21	39	1.6	965	22	AR024329	Sequence 22 from paten	3.50e-01
22	38	1.6	1024	20	BT079414	Bos taurus common all	3.50e-01
23	38	1.6	1285	24	ADPGIP	A. deliiosa pigip mRNA	3.50e-01
24	38	1.6	1425	22	AR005195	Sequence 1 from paten	3.50e-01
25	38	1.6	1425	22	181228	Sequence 1 from paten	3.50e-01
26	38	1.6	1425	22	181228	Sequence 1 from paten	3.50e-01
27	38	1.6	1451	27	AF013988	Sequence 1 from paten	3.50e-01
28	39	1.6	1650	20	SS049435	Homo sapiens serine pr	3.50e-01
29	38	1.6	1799	22	AF024082	Sus scrofa growth horm	1.69e-01
30	38	1.6	2399	22	AF024082	synthetic gene for pla	3.50e-01
31	38	1.6	2573	19	DM049724	Drosophila melanogaste	3.50e-01
32	38	1.6	2775	21	PAFR0AT	Pseudoplatonectes ame	3.50e-01
33	38	1.6	2862	29	S67465	angiotensin II type 2	3.50e-01
34	38	1.6	2862	22	I26126	Sequence 1 from paten	3.50e-01
35	38	1.6	3433	27	HS338H10	Human DNA sequence fro	1.69e-01
36	39	1.6	33760	26	HSNFG9	Human DNA sequence fro	1.69e-01
37	38	1.6	35730	27	HSN74G7	Human DNA sequence fro	1.69e-01
38	39	1.6	74371	27	AC005369	Human DNA sequence fro	1.69e-01
39	39	1.6	151840	18	HS919B11	Homo sapiens chromosom	1.69e-01
40	39	1.6	159144	18	AC005849	Human DNA sequence ***	1.69e-01
41	39	1.6	183827	28	AC005899	*** SEQUENCING IN PROG	1.69e-01
42	39	1.6	212175	18	HS75N14	Homo sapiens chromosom	1.69e-01
43	39	1.6	130027	27	AC004982	Human DNA sequence ***	1.69e-01
44	37	1.5	189666	18	AC000052	Homo sapiens PAC clone	7.20e-01
45	37	1.5	220000	18	AC004480	*** SEQUENCING IN PROG	7.20e-01

ALIGNMENTS

RESULT	LOCUS	DEFINITION	ACCESSION	NID	KEYWORDS	SOURCE
1	G22461	human STR WI-30695.	G22461	91342787	STR sequence; primer; sequence tagged site.	human STRs derived from sequences in dbEST and the Unigene collection.

ORGANISM
 Homo sapiens
 Eukaryote; Mitochondrial eukaryotes; Metazoa; Chordata;
 Vertebrates; Eutheria; Primates; Catarrhini; Homidae; Homo.
 1 (bases 1 to 271)

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 Whitehead Institute/MIT Center for Genome Research, Physically
 Mapped STRs
 Unpublished (1995)

Contact: Thomas Hudson
 Whitehead Institute/MIT Center for Genome Research
 Whitehead Institute for Biomedical Research
 9 Cambridge Center, Cambridge MA 02142 USA
 Tel: 617 252 1900
 Fax: 617 252 1902
 Email: thudson@genome.wi.mit.edu

Primer A: CCCGAGTGTAGGAAATCTGT
 Primer B: CTGGAGATCAAGGCGCT
 STR size: 112
 PCR Profile:

Preheat:
 Denaturation:
 Annealing: 56 degrees C
 Polymerization:
 PCR Cycles: 35
 Thermal Cycler:
 Protocol:
 Template: 10 ng

Primer: each 5 pm
 dNTPs: each 4 mM
 Taq Polymerase: 0.025 units/ul
 Total Vol: 20 ul

Buffer:
 MgCl₂: 1.5 mM
 KCl: 50 mM
 Tris-HCl: 10 mM
 pH: 9.3

FEATURES
 Source Derived from dbEST (genbank accession H18836).
 Location/Qualifiers
 1. 271
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /map="349.5 CR from top of Chr19 linkage group"

STS
 primer_bind 32..143
 primer_bind 32..143
 Complement(126..143)
 BASE COUNT 50 a 64 c 86 g 67 t 4 others
 ORIGIN

Query Match
 Best Local Similarity 97.8%; Score 258; DB 31; Length 271;
 Matches 265; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

Db 1 TTTTTCCTTCCTGAATGGCTTATTTCCCACTGAGGAATCTGTAGTGTGA 60
 CP 2361 TTTTTCCTTCCTGAATGGCTTATTTCCCACTGAGGAATCTGTAGTGTGA 2302
 Db 61 GGGCTGTGTGCTCCTGTACCCCAAGGACCTGACCTCCAGATGATGATGAGGCGG 120
 CP 2301 GGGCTGTGTGCTCCTGTACCCCAAGGACCTGACCTCCAGATGATGATGAGGCGG 2242
 Db 121 GATCCAGGCTTGTATCTCCAGAGTGGAGTGGTCTGTGGAGAGCCGGGCTGTG 180
 CP 2241 GATCCAGGCTTGTATCTCCAGAGTGGAGTGGTCTGTGGAGAGCCGGGCTGTG 2182
 Db 181 GTTGTAAAGTACACTCCCAAGGAGTGTCTCTGAGAGTGGCCAGATGATGAGG 240
 CP 2181 GTTGTAAAGTACACTCCCAAGGAGTGTCTCTGAGAGTGGCCAGATGATGAGG 2122
 Db 241 CACCTNAAGGACCAAGGCGCAGAGTTCGG 271
 CP 2121 CACCTCAAGGACCAAGGCGCAGAGTTCGG 2092

RESULT 2
 LOCUS AB001335 6220 bp mRNA PRI 28-NOV-1998
 DEFINITION Homo sapiens mRNA, complete cds.
 ACCESSION AB001335
 NID G3928755
 KEYWORDS KNP3
 SOURCE Homo sapiens 20-26weeks fetus brain tissue, lib:human fetal brain
 5'-STRECH Plus cDNA library (CLONTECH) and Human Brain, caudate
 nucleus 5'-STRECH cDNA library (CLONTECH) cDNA to mRNA.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
 Primates; Catarrhini; Hominoidea; Homo.
 1 (sites)
 Nagamine, K., Kudo, J., Minoshima, S., Kawasaki, K., Asakawa, S.,
 Ito, F. and Shimizu, N.
 Molecular cloning of a novel putative Ca²⁺ channel protein (TRPC7)
 highly expressed in brain
 Genomics 54 (1), 124-131 (1998)
 2 (bases 1 to 6220)
 Shimizu, N.
 Direct Submission
 Submitted (28-FEB-1997) to the DDBJ/EMBL/Genbank databases.
 Nobuyoshi Shimizu, Keio University School of Medicine, Department
 of Molecular Biology, 35 Shinanomachi, Shinjuku-ku, Tokyo 160,

Japan (E-mail: shimizun@med.keio.ac.jp, Tel: 03-3351-2370,
 Fax: 03-3351-2370)
 Location/Qualifiers
 1. 6220
 /organism="Homo sapiens"
 /note="7 Caucasian"
 /db_xref="taxon:9606"
 /chromosome="21"
 /dev_stage="20-26weeks fetus"
 /map="21q22.3"
 /tissue="lib+human fetal brain 5'-STRECH Plus cDNA library
 (CLONTECH) and Human Brain, caudate nucleus 5'-STRECH cDNA
 library (CLONTECH)"
 /tissue_type="brain"
 446..4957
 /gene="KNP3"
 446..4957
 /note="KNP3"
 /note="similar to C.elegans hypothetical protein
 C01H8.1, C03C03.3, C03F4D1.5, similar to trp and
 trp-like proteins"
 /codon_start=1
 /db_xref="PID:d1035682"
 /db_xref="PID:g3928756"
 /translation="MEPSALRKAGESEGEFEGQLPRVTDGMYSNLRSSSLFKS
 RLQCPFGNNDKQESLSWIPENIKKKECVFVESKLSADKRVCCQCYTHQHLPEA
 TKPTFOGTQDPAKHOEMPTDAPGDVPGFGLSOKRYKRVSDITPSVTHLMTQ
 HMGIDVPLRLISYGAANKPMKRLSIFRGLKVAQTGAMITGSGTGVKQV
 GEAVRDSLSYSIEGELITGVATKGVHNRGLIHPTGSEFPAEYILDEGQCNLTC
 LDSNHSRFTLVDDTGHQGYEIPRLTRLEFIESEOTERGVAKIPYCVVEGGP
 GTLTIDNATNGTCCVVEGSGRAVIAQVANLPVSDITISLQKLSFPEMPE
 TFESRIYEMTKKIODIYRRQLTLVREGDGODVDATLQALLKRSQDFHGE
 NMDQLKLVAMNKKVODIYRRQLTLVREGDGODVDATLQALLKRSQDFHGE
 OLKEFTVMDLXYIKENDPSCLFHSKLVQVDEPERPACAPAPLQMHVQVNR
 ELDDFTQPLPRKRNDRLLLPVHVKNLVQVSLRSKRSCHVTFFPIRD
 LLIAIVQNRRELGIIMASODCIAMALCSKLKESKEEDDSSEMLAEFY
 EHRAGIVTECYRDEERAKLLRVSEANGKTCQLDLAENAKMKRSIGGICAT
 KYMGQSLSYDNGRLVRLKLAFLRLVGLISPERKLDVGTAPAAARFFAPAVV
 FHLNLSYAFILCPAYLVAVNDQPVSKSCATYLLFSLVCEEMQLYDDEGCL
 MKRALFSPFNMKLDGAILLPVAGLTLCLIPATLPVTLISLPIELCLRMHIF
 TISTTLPIKTIIVAKMKDVEFLFLAVVWSGVAKQAILINERVMDLFRGAVY
 HSITITGIGIIPGYIDGVNFPNHCSPGTDIPYKPCPSDATORAPAPFELVLLC
 LYLTFILILINLIANFNTPFOOVDDQINKFQALINERVMDLFRGAVY
 HLOLFIRVYAKTPAKRHKLKNKLENEBAALISWEIYAKENTLROKOROPRILS
 KIEDISNKVDAMVDLIDPLKRSSEGLASIEEDVQATARLHIVTLAASGS
 SEADVPLAOKAABEPDAEPGKTEEPQDSTHVANRLLYPNCVPTFPVNERV
 PWTEFTLIDPPTIARKDAAMADPGLTLEPLSTIOYNVVDGLRBRSGFPGYVQ
 AGLPNMGRTGRLGRGSLSCFSPNHTLYPMTVRMRNEDGACRSIKKMLEVLYV
 LPLSEHMLPGSREPEMLPKRLRLRQEHNPSPENLKCMEYVKCYNDPRTD
 NATEYVAVSVHFQODNDVLELNLSLHNCDSGASIRQVYDRIPLVIANHKLQK
 AAEFGAHY"

FEATURES
 Source

gene
 CDS

BASE COUNT 1248 a 1938 c 1882 g 1152 t
 ORIGIN

Query Match
 Best Local Similarity 4.9%; Score 118; DB 28; Length 6220;
 Matches 423; Conservative 0; Mismatches 269; Indels 6; Gaps 4;

Db 3092 TACTTCAGTACTCTGGAATTAAGTCTGAGCGGCGGCAATCTTCTGTCGAGG 3091
 QY 945 TACTTCAGTACTCTGGAATTAAGTCTGAGCGGCGGCAATCTTCTGTCGAGG 3091
 Db 3092 CTGACCTGACGACATCGCGGAGAGCTTACCCGCGGCGGACATCTCTCTGAGC 3151
 QY 1005 GTGGGCTGCGGCGGAGAGCTTACCCGCGGCGGACATCTCTCTGAGC 3151
 Db 3152 TTCACTCTGTCGCTCGGCTGACATGACATTTTACATGATGAGAGCTGGGCCC 3211
 QY 1065 TTCACTCTGTCGCTCGGCTGACATGACATTTTACATGATGAGAGCTGGGCCC 3211
 Db 3212 AGATATCATCTTGAAGGAGATGAGAGAGCTCTTCTCTCTCTCTCTGTCGCT 3271
 QY 1125 AGATGTCATCTGAGCAAGATGAGAGAGCTCTTCTCTCTCTCTCTCTCTGTCGCT 1184

REFERENCE 3 (bases 1 to 155074)
 AUTHORS Evans, G.A., Athanasiou, M., Aguayo, P., Armstrong, D., Basil, M., Buettner, J., Butler, C., Card, P., Desaliboat, F., Dunn, J., English, C., Ethridge, S., Garner, H.R., Gee, Y., Gordon, M., Gotway, G., Grant, O., Hahner, L., Joslin, J., Lewis, E., Loo, H., Loo, K.N., Major, T., McFarland, J., Newton, J., Osborne-Lawrence, S., Schageman, J., Schultz, R.A., Stimson, S., Waller, K. and Ward, T.
 TITLE Direct Submission
 JOURNAL Submitted (30-SEP-1998) Genome Science & Technology Center, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd., Dallas, TX 75235-8591, USA
 On Oct 1, 1998 this sequence version replaced gi:3264564.
 COMMENT Further information regarding the map of this region or annotation of PDJ915f1 can be found at
<http://gestec.swmed.edu/chromoso.htm>
 IMPORTANT: This submission contains the entire insert of clone PDJ915f1. PDJ915f1 comes from the RPCI-3 PAC library constructed at the Roswell Park Cancer Institute by the Plaster de Jong group. This clone has been finished according to strict quality criteria and attempts have been made to resolve all base calling problems such as compressions and repetitive elements. The expected phred/phrap calculated error/10kb is 0.97. In addition, attempts have been made to assure over 99% of consensus base calls consist of either double-stranded coverage or 2 types of labeling chemistry on one strand.

CHROMOSOMAL LOCUS: This PAC clone comes from the Chromosome 11p15.5 Williams tumor type 2 (WT2) region between ESTs CTSD and IGF2(INS12). This region spanning approximately 2 Mbp is mapped between D1S2701 and D1S1145.

MARKER CONFIRMATION: STS WI-9248.

FEATURES

SOURCE

1. 155074
 Location/Qualifiers

repeat_region	/organism="Homo sapiens"	complement(68934..69048)
repeat_region	/db_xref="taxon:9606"	complement(68934..69048)
repeat_region	/rpt_family="Alu"	complement(68934..69048)
repeat_region	complement(12602..12763)	complement(68934..69048)
repeat_region	/rpt_family="MER4"	complement(68934..69048)
repeat_region	complement(13184..13415)	complement(68934..69048)
repeat_region	/rpt_family="Alu"	complement(68934..69048)
repeat_region	complement(26921..27026)	complement(68934..69048)
repeat_region	/rpt_family="Alu"	complement(68934..69048)
repeat_region	32140..32471	complement(68934..69048)
repeat_region	/rpt_family="Alu"	complement(68934..69048)
repeat_region	34325..34628	complement(68934..69048)
repeat_region	/rpt_family="Alu"	complement(68934..69048)
repeat_region	39386..39469	complement(68934..69048)
repeat_region	/rpt_family="MLT1"	complement(68934..69048)
repeat_region	39630..39771	complement(68934..69048)
repeat_region	/rpt_family="MLT1"	complement(68934..69048)
repeat_region	44752..44912	complement(68934..69048)
repeat_region	/rpt_family="THE1"	complement(68934..69048)
repeat_region	45325..45410	complement(68934..69048)
repeat_region	/rpt_family="THE1"	complement(68934..69048)
repeat_region	complement(49939..50077)	complement(68934..69048)
repeat_region	/rpt_family="Alu"	complement(68934..69048)
repeat_region	52802..55183	complement(68934..69048)
repeat_region	/rpt_family="L1"	complement(68934..69048)
repeat_region	53582..53663	complement(68934..69048)
repeat_region	/rpt_family="Alu"	complement(68934..69048)
repeat_region	complement(56684..56986)	complement(68934..69048)
repeat_region	/rpt_family="Alu"	complement(68934..69048)
repeat_region	complement(57763..57988)	complement(68934..69048)
repeat_region	/rpt_family="L1"	complement(68934..69048)
repeat_region	complement(59538..59895)	complement(68934..69048)
repeat_region	/rpt_family="TAR1"	complement(68934..69048)
repeat_region	complement(66135..66214)	complement(68934..69048)
repeat_region	/rpt_family="MER1"	complement(68934..69048)
repeat_region	complement(66228..66434)	complement(68934..69048)
repeat_region	/rpt_family="MER1"	complement(68934..69048)
repeat_region	68089..68377	complement(68934..69048)
repeat_region	/rpt_family="Alu"	complement(68934..69048)
repeat_region	complement(68934..69048)	complement(68934..69048)
repeat_region	/rpt_family="L1"	complement(68934..69048)
repeat_region	complement(70090..70376)	complement(68934..69048)
repeat_region	/rpt_family="Alu"	complement(68934..69048)
repeat_region	71058..71304	complement(68934..69048)
repeat_region	/rpt_family="Alu"	complement(68934..69048)
repeat_region	71629..71964	complement(68934..69048)
repeat_region	/rpt_family="Alu"	complement(68934..69048)
repeat_region	complement(80143..80239)	complement(68934..69048)
repeat_region	/rpt_family="Alu"	complement(68934..69048)
repeat_region	complement(83167..83415)	complement(68934..69048)
repeat_region	/rpt_family="Alu"	complement(68934..69048)
repeat_region	complement(84253..84559)	complement(68934..69048)
repeat_region	/rpt_family="Alu"	complement(68934..69048)
repeat_region	complement(84663..85133)	complement(68934..69048)
repeat_region	/rpt_family="Alu"	complement(68934..69048)
repeat_region	90644..90990	complement(68934..69048)
repeat_region	/rpt_family="THE1"	complement(68934..69048)
repeat_region	complement(91700..91966)	complement(68934..69048)
repeat_region	/rpt_family="Alu"	complement(68934..69048)
repeat_region	92586..92737	complement(68934..69048)
repeat_region	/rpt_family="Alu"	complement(68934..69048)
repeat_region	92944..93169	complement(68934..69048)
repeat_region	/rpt_family="MER42"	complement(68934..69048)
repeat_region	complement(93315..93666)	complement(68934..69048)
repeat_region	/rpt_family="THE1"	complement(68934..69048)
repeat_region	complement(93707..95255)	complement(68934..69048)
repeat_region	/rpt_family="MSR4"	complement(68934..69048)
repeat_region	complement(95260..95652)	complement(68934..69048)
repeat_region	/rpt_family="THE1"	complement(68934..69048)
repeat_region	95968..96013	complement(68934..69048)
repeat_region	/rpt_family="MER42"	complement(68934..69048)
repeat_region	96486..97004	complement(68934..69048)
repeat_region	/rpt_family="MER1"	complement(68934..69048)
repeat_region	complement(98550..98704)	complement(68934..69048)
repeat_region	/rpt_family="MER42"	complement(68934..69048)
repeat_region	complement(99173..99442)	complement(68934..69048)
repeat_region	/rpt_family="Alu"	complement(68934..69048)
repeat_region	complement(99456..99916)	complement(68934..69048)
repeat_region	/rpt_family="MER42"	complement(68934..69048)
repeat_region	complement(10086..103880)	complement(68934..69048)
repeat_region	/rpt_family="L1"	complement(68934..69048)
repeat_region	complement(101541..101608)	complement(68934..69048)
repeat_region	/rpt_family="THE1"	complement(68934..69048)
repeat_region	complement(105454..105672)	complement(68934..69048)
repeat_region	/rpt_family="MER43"	complement(68934..69048)
repeat_region	complement(105673..105983)	complement(68934..69048)
repeat_region	/rpt_family="Alu"	complement(68934..69048)
repeat_region	complement(105997..106047)	complement(68934..69048)
repeat_region	/rpt_family="MER43"	complement(68934..69048)
repeat_region	complement(107379..114158)	complement(68934..69048)
repeat_region	/rpt_family="L1"	complement(68934..69048)
repeat_region	111143..111419	complement(68934..69048)
repeat_region	/rpt_family="Alu"	complement(68934..69048)
repeat_region	complement(11213..112455)	complement(68934..69048)
repeat_region	/rpt_family="MER25"	complement(68934..69048)
repeat_region	115302..116008	complement(68934..69048)
repeat_region	/rpt_family="L1"	complement(68934..69048)
repeat_region	116348..119279	complement(68934..69048)
repeat_region	/rpt_family="L1"	complement(68934..69048)
repeat_region	119474..119773	complement(68934..69048)
repeat_region	/rpt_family="Alu"	complement(68934..69048)
repeat_region	125016..125597	complement(68934..69048)
repeat_region	/rpt_family="Alu"	complement(68934..69048)
repeat_region	129543..139819	complement(68934..69048)
repeat_region	/rpt_family="Alu"	complement(68934..69048)
repeat_region	complement(130188..130341)	complement(68934..69048)
repeat_region	/rpt_family="L1"	complement(68934..69048)
repeat_region	134446..134618	complement(68934..69048)
repeat_region	/rpt_family="Alu"	complement(68934..69048)
repeat_region	complement(138188..138488)	complement(68934..69048)

between the conligs are also unknown; these gaps are presented

1.98; Score 45; DB 18; Length 193381;

1.98; Score 45; DB 18; Length 193381;

REFERENCE	AUTHORS	TITLE
1 (bases 1 to 2325)	Parker, E.M., Izazarelli, D.G., Nowak, H.P., Mahle, C.D., Iben, L.G., Wang, J., and Goldstein, M.E.	Cloning and characterization of the rat GALRI galanin receptor from Rln4B insulinoma cells
JOURNAL	Brain Res.	MOL. Brain Res. 34 (2), 179-189 (1995)
MEDLINE	96363004	
REFERENCE	2 (bases 1 to 2325)	
AUTHORS	Parker, E.M.	
TITLE	Direct Submisslon	
JOURNAL	Submitted (25-JUN-1995)	Eric M. Parker, CNS Biology, Dept. # 404, Bristol-Myers Squibb Company, 5 Research Parkway, Wallingford, CT 06492, USA

```
location/qualifiers
1. .2865
/organism="Homo sapiens"
/db_xref="taxon:9606"
```

Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homini

REFERENCE 1 (pages 1 to 3519)

AUTHORS

Martin, C.H., Aicardi, T.T., Bondoc, M.M., Chang, A., Citz, P.A., Davis, C.A., Doyle, C.M., Eliscovich, C.L., Farrant, D.E., Gunning, K.M., Houston, K.A., Jaklevic, C.M.A., Kachner, K.E., Kim, K., Kim, S.F., Ko, C.L., Lewis, K.D., Li, M., Lindquist, R.J., Lomtan, M.A., Lustre, V.M., Mechuris, M.U., Mayeda, C.A., Mignuel, T.M., Miller, C.A., Mok, M.S., Pacלב, J.M., Patel, S.G., Santos, R.F., Subramanian, S., Wan, K.H., Wittelev, K., Yee, A., Yeh, R.T., Yu, C. and Palazzolo, M.J. Sequencing of human chromosome 5q (unpublished, 1996).

TITLE
JOURNAL
REFERENCE
AUTHORS

2 (bases 1 to 3519)